Labor Market Segregation and the Wage Differential between Resident and Migrant Workers in China†

Xiangdong Wei* and Ruosi Lu*

August 2007

Abstract

This paper looks at the effect of industrial and occupational segregation on the wage differential between resident and migrant workers in China. It extends the work of Meng and Zhang (2001) by considering the possible employment segregation of resident and migrant workers by both industry and occupation. We contend that industry segregation is at least as important as occupational segregation for Chinese migrant workers, as most migrant workers in China have come from the countryside to fuel the booming labor-intensive manufacturing and construction industries in the cites. Due to the *hukou* policy (a household registration system) in China, migrant workers normally face more constraints in searching for jobs in other sectors. Our empirical study confirms that the proportion of the resident-migrant worker wage differential that is explained by industrial segregation is much larger than that explained by occupational segregation. Taking both industrial and occupational segregation into account explains the substantial wage differential between resident and migrant workers, which indicates the influence of industrial and occupational barriers on the wage differential in China.

Keywords: Industry, occupation, segregation, wage differential, migrant workers.

JEL Code: J71, J31, O15, P25

†We thank David Drury and Marie Arneburg from the Fafo Institute for providing us with the SOMM data. We are responsible for any remaining errors. We are also grateful for the helpful comments provided by Junsen Zhang and the participants of the 2007 CES Changsha Conference. The usually disclaimer applies.

*Department of Economics, Lingnan University, Tuen Mun, Hong Kong. E-mail: xdwei@ln.edu.hk. Corresponding Author: Xiangdong Wei.

Introduction

After 30 years of economic reform, China has been transformed from a predominantly agricultural economy into an industrial powerhouse. Its urban population has increased from about 12% of the total population in the late 1970s to 40% in 2006 (China Statistical Year Book), a result not only of the rapid rate of urbanization but also of a dramatic increase in the rural to urban migration rate. It is now estimated that there are approximately 0.18 billion rural migrant workers working in the cities who account for more than 50% of the labor force in China's industrial sector[1].

Despite such a fast rate of industrialization and urbanization, income inequality in China as measured both in terms of coastal versus inland and urban versus rural workers has been widening since the mid-1980s. According to the estimation of Lin, Wang, and Zhao (2004), the coast-inland income ratio rose from 1.31 in 1985 to 1.65 in 2000, and the urban-rural income ratio rose from 1.82 in 1985 to 2.42 in 2000. Such a rise in inequality may appear at odds with conventional economic theory, which predicts factor price equalization accompanied by the movement of labor from the countryside to the city or from agricultural activities to industrial activities. As has been repeatedly shown by previous researchers in this area,[2] an important factor that contributes to this anomaly is the special household registration system, or *hukou* system in China. In brief, the household registration system specifies that individuals must register with the local authority at their place of birth, and only registered local people are entitled to the various welfare programs provided by the local government, which include housing, education, health care, unemployment benefits, and income support. Various unemployment reduction policies are also targeted at local residents only. This system greatly disadvantages migrant workers, and helps to create a two-tier labor market in the cities. For example, according to recent surveys (see footnote 1 for the source), rural migrant workers are highly concentrated in low-end, labor-intensive jobs (occupying 57.6% of the industrial sector jobs and 80% of the labor-intensive processing industry jobs) in non-state owned sectors. Only a third of rural migrant workers have signed a formal labor contract with their employers, and their average wage is only half that of resident urban workers. Many migrant workers are not paid on time and have to work extremely long hours without overtime pay.

The status and treatment of these rural migrant workers is so alarming that it is now high on the policy agenda of the central government.[3] However, the issue of wage inequality between urban resident and rural migrant workers is complicated by the fact that the average educational level of rural migrant workers is also significantly lower than that of their urban counterparts. Any government policies that aim to reduce the wage inequality between these two groups must carefully distinguish the wage differential that is created by institutional or discriminative forces from that generated by normal market forces. Thus far, there has been a lack of empirical work to estimate the exact impact of potential labor market segregation on wage inequality between urban and rural migrant workers. The work of Meng and Zhang (2001) is pioneering in this respect.

Meng and Zhang (2001) utilize the data from two surveys conducted in 1999 for Shanghai and the methodology developed by Brown *et al.* (1980) to decompose wage differential into the productivity effect, occupational segregation effect, and unexplained or discrimination effect. Their decomposition shows that occupational segregation only contributes roughly 4.85% of the total wage differential, whereas intra-occupational factors account for 82%. They quote this as a rather surprising finding, yet careful observation of the labor market segregation that rural migrant workers in China face would give rise to the conclusion that industry segregation is at least as important as occupational segregation. We offer three explanations as to why this may be the case. First, rural migrants were first allowed to migrate to cities because certain industries in the cities faced labor shortages due to the compound effect of the fast expansion of the labor-intensive industrial sector and the improvement in living standards in the cities (with the result that urban resident workers became unwilling to work for jobs with low pay and poor working conditions). Second, the existence of the household registration system means that rural migrant workers are entitled to less welfare and other government support, and hence their ability to search for and move to better jobs is much less than that of their urban counterparts. However, it may be relatively easier for them to move up the occupational ladder once employed. Third, due to information costs and other barriers to entry, most rural migrant workers rely on their informal social network to obtain jobs (Zhao, 2001). The usual practice is that pioneers go to the city and find jobs first and then introduce their friends and relatives to the same jobs in the same industry. Thus, given the special circumstances in China, the labor market segregation for rural migrants may take the form of both occupational crowding and industry crowding,[4] as reflected by the figures in Table 1.

The current study makes its contribution in this area. More specifically, we extend the Brown et al (1980) approach by taking into account both the industry and occupational segregation effects. In so doing, we separate workers into four groups: blue-collar workers in the industrial sector, white-collar workers in the industrial sector, blue-collar workers in the service sector, and white-collar workers in the service sector. Furthermore, our study utilizes data from a 1998 survey conducted by the Fafo Institute for Applied International Studies in Oslo and the National Research Center for Science and Technology for Development (NRCSTD), which was the first integrated survey of resident and migrant workers covering several major Chinese cities. The survey also collected detailed information on the personal and job characteristics of the respondents, which enables us to better estimate the first-stage multinomial logit models of job attainment. Our results show that 47% of the wage differential is attributable to between-group effects, and 37% is associated with industry and occupational segregation (unexplained between-group effects). This finding confirms that labor market segregation is indeed an important factor that contributes to the wage inequality between urban resident and rural migrant workers.

The remainder of the paper is as follows. We outline our empirical methodology in section 2. Section 3 describes the data and sample characteristics. We present and discuss our empirical findings in section 4, and section 5 concludes the paper.

2. Methodology

The Blinder-Oaxaca (1973) decomposition approach is the most popular method for analyzing the wage differential. The implementation of this approach requires the estimation of a wage determination equation, and the specification of the wage equation is mainly based on the human capital theory. The rationale of this approach is that a wage differential can be decomposed into two parts: one that can be explained by human capital factors and the other that cannot be explained by observed individual productivity characteristics and is hence ascribed to labor market discrimination. However, as discrimination can take the form of both unequal access to jobs and unequal treatment in the same job, this method is unable to distinguish which form contributes more to the observed wage differential.

Brown *et al.* (1980) extend the Blinder-Oaxaca approach by explicitly incorporating the effect of occupational segregation in their decomposition. They assume that there may be unequal access to occupations for minority or disadvantaged workers in the labor market. As long as the factors that determine a worker's choice of occupation are not completely the same as those that influence a worker's wage, the Blinder-Oaxaca type of wage decomposition may be biased. Similarly, if one believes that there is unequal access to industries for rural migrant workers in China, then potential bias may also arise if industry choice is not controlled for. In this paper, we extend the method of Brown *et al.* by explicitly taking into account the choice of both industry and occupation when estimating the wage equation and decomposing wage differentials.

More specifically, let us assume the wages for urban resident and rural migrant workers are determined by the following Mincerian wage equations.

$$\ln w_j^k = X_j^k \beta_j^k + \varepsilon_j^k \qquad k = u,\ r \text{ and } j = 1, \ldots, J, \tag{1}$$

where the superscripts u and r denote urban resident and rural migrant workers, respectively; w is the earnings of a worker in job j (for simplicity, we suppress the individual subscript i in the equation) and j is an industry-occupation indicator; X is a vector of the individual and job characteristic variables that affect a worker's pay; β is the unknown parameter to be estimated; and ε is the error term.

Next, we assume that the individual industry or occupation choice is determined by the following multinomial logit model.

$$p_{nj}^k = prob(y_n^k = j) = \frac{\exp(Z_n^k \gamma_j^k)}{\sum_{h=1}^{J} \exp(Z_n^k \gamma_h^k)} \qquad k = u,\ r;\ n = 1, \ldots, N \text{ and } j = 1, \ldots, J, \tag{2}$$

where p_{nj}^k is the probability that individual n is working in the jth category of the industry-occupation; N is the sample size; J is the total number of industry-occupation categories; Z represents a vector of the exogenous variables that affect labor supply and demand; and γ is its coefficient. As each individual must select one industry-occupation, only *J*–1 sets of coefficients are uniquely defined. We choose to normalize the coefficients for the first industry-occupation category (blue-collar workers in the industrial sector) to be zero.

Following Brown *et al.*, the wage differential of urban residents and rural migrant workers can be decomposed as:

$$\overline{\ln w}^{u} - \overline{\ln w}^{r} = \underbrace{\sum_{j=1}^{J} p_j^r \hat{\beta}_j^u (\overline{X}_j^u - \overline{X}_j^r)}_{WE} + \underbrace{\sum_{j=1}^{J} p_j^r \overline{X}_j^r (\hat{\beta}_j^u - \hat{\beta}_j^r)}_{WU} + \underbrace{\sum_{j=1}^{J} \overline{\ln w}^{u} (p_j^u - \hat{p}_j^r)}_{BE} + \underbrace{\sum_{j=1}^{J} \overline{\ln w}^{u} (\hat{p}_j^r - p_j^r)}_{BU}$$

The upper bars denote the average, and the upper hats denote the predicted values. The term *WE* is the explained wage differential due to differences in personal characteristics; *WU* is the unexplained wage differential due to differences in the coefficients of the estimated resident and migrant wage equations; *BE* is the explained between-industry-occupation wage differential due to differences in qualifications in an industry-occupation group; and *BU* is the unexplained wage differential due to differences in the structures of industry-occupational attainment. Therefore, the foregoing wage decomposition approach takes into account not only the wage discrimination that arises from doing the same job (*WU*), but also the wage discrimination that arises from unequal access to jobs (*BU*). As an extension of Brown *et al.*, we define jobs by both occupation and industry, which allows us to take into account the labor market segregation that arises both from occupation barriers and industry barriers.

The estimation of (1) and (2) is more efficient if we also take into account the fact that there may be unobserved variables that affect both the wage and industry-occupation choices of workers, and thus the error terms of (1) and (2) may be correlated (see Liu *et al.* 2004). To consider this fully, a two-stage selectivity model must be estimated. In the first stage, (2) is estimated, and in the second stage the following wage equation is estimated.

$$\ln w_j^k = X_j^k \beta_j^k + \alpha_j^k \lambda_j^k + e_j^k \,, \qquad (3)$$

where λ_j^k is the estimated inverse Miller's ratio from the first stage of the estimation. We estimate this multinomial selectivity model separately for urban resident and rural migrant workers. To further improve the efficiency of the estimation, (2) and (3) are estimated jointly by the maximum likelihood method using LIMDEP.

3. Data

The study utilizes data from the 1998 Survey of Occupational Mobility and Migration collected by the Fafo Institute for Applied International Studies in Oslo and the National Research Center for Science and Technology for Development (NRCSTD) in Beijing. The former is an independent non-profit making research institute, and the latter is a branch institute of the Ministry of Science and Technology of China. The survey was carried out in three Chinese cities: Beijing, Wuxi, and Zhuhai (see Appendix 1 for the locations of these three cities). As pointed out by the survey organizers, these three cities were not randomly chosen, but were selected to "explore the effects of the transition in cities of different scale, region, and with different economic profiles" (Drury and Arneberg, 2001, p. 4). Beijing, as the capital of China, is dominated by the public service sector and large state enterprises. Its labor market is more diversified due to its size, but is less open compared with the market in the other two cities. Wuxi is a flourishing industrial city near Shanghai in the Yangtze River Delta area of Jiangsu province, and has followed a model of development that is based on collective and township-village enterprises. It is also a city chosen by the central government to test its new state enterprise reform policies. Zhuhai, as one of the earliest special economic zones in China, is dominated by joint-venture and foreign investment firms, and has the most developed labor market of the three. All three cities have absorbed a large inflow of rural immigrant workers due to their fast economic development, and together give a good representation of well-developed cities in China.

The survey samples were selected randomly in the three cities, with selection being carried out separately for local and migrant workers. A two-stage cluster sampling approach was used to obtain the local resident sample. In the first stage, a random stratification sample of Residential Committees[5] was selected based on location and size. In the second stage, a random sample of households within each Residential Committee was chosen to take part in a household survey. A separate group of clusters based on the neighborhood-level police stations was selected to obtain the migrant sample, and a random sample of migrant households was then selected to take part in the survey.

The survey questionnaire had two parts. The first part was conducted at the household level and aimed to collect information about all of the household members, and the second part was for a randomly selected household member aged 16 or above. As migrants are minorities in the city, these individuals had to be over-sampled to achieve a suitably sized migrant sample. The detailed working history information over the previous five years (1994-1998) was collected from the selected individuals. Together, the two parts of the survey obtained both detailed household information and detailed working history information of an adult member within each household. We use this data for our study because the survey was the first major integrated survey of residents and migrants in cities in China.

The target sample size of the survey was 7,835 households, and the final completed sample contained 7,326 households. The sample sizes for Beijing, Wuxi, and Zhuhai were 2,446, 2,437, and 2,443, respectively. Due to missing information, our final sample contains 3,886 workers, of which 1,682 are migrant workers and 2,204 resident workers.

As mentioned, we define four industry-occupation groups: (1) blue-collar workers in the industrial sector; (2) white-collar workers in the industrial sector; (3) blue-collar workers in the service sector; and (4) white-collar workers in the service sector. The classification of industries into either the industrial sector or the service sector closely matches the standard broad definition of sectors used by the China Statistical Bureau, i.e. the primary sector (agriculture), the secondary sector (manufacturing, construction, and public utilities), and the tertiary sector (the service sector). It should be noted that our use of the term "industrial sector" covers the whole secondary sector, not just the manufacturing industry. This broad definition of industry and occupation groups is used mainly because of the sample size consideration, as we do not want to end up with industry or occupation cells that contain only a handful of observations.

The data in Table 1 show that rural migrant workers are far more concentrated in the blue-collar and industrial sector group (64.21%) and far less represented in the white-collar and service sector (3.86%) than their urban counterparts (the corresponding figures for the latter are 28.77% and 33.08%, respectively). Furthermore, the mean log hourly wage for blue-collar workers in the industrial sector is the lowest and that for white-collar workers in the service sector the highest for both rural migrant and urban resident workers. These raw data indicate that our classification of industry-occupation groups captures the main thrust of labor market segregation in China well.

For the multinomial logit model of industry-occupation selection, the vector of the independent variables Z should include those variables that affect both the labor supply and demand for a job. The factors influencing an individual's supply of labor are wealth, preferences, and job search costs, whereas the factors that determine the demand for labor are mainly those affecting an individual's productivity. We thus include in Z years of work experience, years of schooling, a dummy variable for gender, a dummy variable for marital status, dummy variables that indicate whether the respondent's father is a party cadre or is self-employed (owns a business), a dummy variable for working in a state-owned enterprise, a dummy variable for getting the job through state allocation, and dummy variables for the three cities. The human capital variables that are included are meant to control for both an individual's productivity and wealth, whereas the family background variables are proxies for individual preferences. The dummy variable for getting a job through state allocation and the city dummies are related to the job search costs. We believe that the family background variables and the dummy for whether the respondent obtained a job through state allocation are the main identification variables that should be included in the industry-occupation choice equation, but not the wage equation.

The specification of our wage equation follows the usual human capital theory. It includes a dummy for gender, years of work experience and its square, years of schooling, a dummy that indicates whether an individual is a party cadre, a dummy for working in a state-owned enterprise, and two city dummies.[6]

The variable definition and descriptive statistics are reported in Table 2. As shown in the table, on average the migrants earn roughly 38% less than the residents, and also tend to have less human (years of schooling) and political (Party cadre) capital. The fewer years of work experience and lower probability of being married among the migrants are mainly due to their younger age profile compared with the residents. Furthermore, the migrants are less likely to have a father who is a Party cadre, less likely to work in a state-owned enterprise, and less likely to have gotten their job through state allocation. However, they are more likely to have a self-employed father.

The most interesting difference between the migrant and resident workers in this study is obviously the wage differential. In what follows, we examine the extent to which this differential can be explained by potential labor market segregation.

4. Empirical results

4.1. Multinomial logit model of industry-occupation attainment

As mentioned, our first step in estimating the selectivity model is to run a multinomial logit model of industry-occupation attainment separately for residents and migrants. The results are presented in Table 3, which only reports the estimated coefficients for three of the industry-occupation groups, as the coefficient for the base group – blue-collar workers in the industrial sector – is set to zero. Consequently, all of the coefficients should be interpreted by comparison with the base group. Years of schooling has a positive and significant effect on an individual's chances of entering the service sector and holding a white-collar job. In contrast, work experience seems to have no significant impact on industry-occupation attainment, as is the case with marital status. Being a party cadre significantly increases an individual's chances of working in the service sector or as a white-collar worker. This is also true if an individual has a father who is a Party cadre, which indicates that political capital is important in job attainment in China. State-owned enterprises are more likely to provide white-collar jobs in the service sector, but this may simply reflect the fact that the government still monopolizes top-end service activities in China. Individuals with jobs allocated by the government are more likely to be blue-collar workers in the manufacturing industry, as are residents of Beijing or Wuxi.

In Table 5, we report the actual and predicted probability of job attainment for the various occupation-industry groups. We can see that the predicted proportion of rural migrant workers working as blue-collar workers in the industrial sector is far below the actual proportion, whereas the predicted proportion of rural migrant workers working as white-collar workers in the service sector is far higher than the actual proportion. In contrast, the predicted proportion of urban resident workers working as white-collar workers in the service sector is much less than the actual proportion. In other words, other things being equal, rural migrant workers are more likely to end up as blue-collar workers in the industrial sector, and urban resident workers are more likely to be found in white-collar jobs in the service sector. This indicates that there is indeed labor market segregation in China for urban resident and rural migrant workers.

Finally, we find that the dummy for whether an individual's father is a cadre and whether the respondent obtained his or her current job through state allocation are significant in most of the multinomial logit models. The dummy for whether an individual's father owns a business also appears to be significant in one model (blue-collar workers in the service industry). Together, these results show that our identification variables for the selectivity model indeed play a significant role in the job attainment equations.

4.2. Selectivity-corrected industry-occupation specific earnings functions

The selectivity-corrected wage equations for the various occupation-industry groups are reported in Table 4. All of the human capital variables have their expected signs. The rates of return to education range from 3% to 9%, and are higher for migrants in the service sector. Work experience always enters positively, but is only significant in three out of the eight regressions. This is consistent with previous findings for China and is probably due to the fact that older workers had a much smaller wage increment before the reform era. Interestingly, the political capital indicator – the dummy for being a party cadre – is only positive and significant for the resident white-collar industrial sector workers, resident blue-collar service sector workers, and migrant white-collar service sector workers. Finally, state-owned enterprises tend to pay workers – both residents and migrants – more in the industrial sector, but almost the opposite is true in the service sector.

We run the specification with the two family background dummies and the state job allocation dummy included in the wage equation, but none is significant. We therefore assert that our selectivity model meets the identification condition.

4.3. Wage decomposition

To perform the wage decomposition outlined in the methodology section, we need to estimate the nondiscriminatory industry-occupational distribution $\hat{p}^r$ for the migrants.

This is undertaken by substituting the observed migrant characteristics into the estimated resident industry-occupation attainment model. The estimation of the other terms in the wage decomposition formula is straightforward. We present the estimates in Table 5, from which we can see that the predicted industry-occupational distributions for the migrants differ significantly from the observed distributions. As expected, the estimated probability of migrants being blue-collar workers in the industrial sector is far lower than the actual probability, and the estimated probability of migrants being white-collar workers in the service sector is far larger than the actual probability.

In Table 5, the estimated within group explained wage differential (WE) is 0.1042, or 27% of the total wage differential. The estimated within group unexplained wage differential (WU) is almost the same, at 0.1004, or 26% of the total wage differential. This may be viewed as the upper bound of the on-the-job wage discrimination against migrant workers. The estimated between-group explained wage differential (BE) is only 0.0389, or 10%, which is much smaller than the estimated between-group unexplained wage differential (BU) of 0.1416, or 37%. This latter figure shows that labor market segregation or unequal access to industry-occupation contributes as much as 37% of the observed wage differential between resident and migrant workers.

Our estimated results here are significantly different from those obtained by Meng and Zhang (2001). First, our estimated between-sector wage differential represents 47% (10% + 37%) of the total wage differential, whereas the estimation of Meng and Zhang, which is based on occupational segregation alone, is only 12%. The second, and perhaps most important, difference is that our estimated labor market segregation effect accounts for as much as 37% of the total wage differential, which is in sharp contrast to the 4.85% figure estimated by Meng and Zhang. We ascribe the significant difference between the two studies to our use of both industry and occupation to define labor market segregation. To further support this claim, we also experiment with wage decomposition using either occupation or industry.

In Table 6, we present the results for a multinomial logit selectivity-adjusted wage model that uses occupation alone as the labor market segregation measure (blue- versus white-collar workers), and the estimated total between-occupation wage differential drops to 34%. More interestingly, the estimated between-occupation unexplained wage differential is only 0.06%, which is even smaller than the 4.85% estimated by Meng and Zhang. This indicates that the impact of labor market segregation on the wage differential due to occupational segregation is negligible.

In Table 7, we estimate our multinomial logit selectivity-adjusted wage model using industry alone as the labor market segregation measure (industrial versus service sector). Although the estimated total between-sector wage differential is still small and accounts for only 19% of the total wage differential, the estimated between-sector unexplained wage differential jumps up to 16%, indicating a significant labor market segregation effect on the wage differential due to unequal access to industries.

5. Conclusions

To conclude, we extend the wage decomposition framework of Brown *et al.* by taking into account potential segregation in both industry and occupation. We apply this framework to decompose the wage differential between urban resident and rural migrant workers in China using data from the 1998 Survey of Occupational Mobility and Migration collected by the Fafo Institute for Applied International Studies in Oslo and the National Research Center for Science and Technology for Development (NRCSTD). In contrast to earlier findings by Meng and Zhang (2001), we find labor market segregation in China as measured by both industry and occupational segregation to contribute significantly (up to 37%) to the observed wage differential between these two groups of workers.

We also perform our analyses using industry or occupation alone as the measure for labor market segregation, and find that the main effect of labor market segregation on the wage differential comes from industry, rather than occupational, segregation. We believe that this finding fits intuition well. Because of the household registration system in China, rural migrant workers face more difficulties in searching for jobs in cities and hence are less likely to switch industries. However, it may be less difficult for them to climb up the occupational ladder within the same firm or industry.

Our findings provide useful policy implications for the government. They show that the household registration system is a major obstacle for migrant workers in gaining equal access to good jobs, and thereby lowers their wages. If the government intends to reduce the earnings gap between resident and migrant workers, then it must abandon this policy, but it should also be abandoned because it increases labor market rigidity and reduces efficiency.

Notes

[1]See the reports by Xinhua News Agency on January 18 and February 24, 2006 (www.xinhuanet.com).

[2]For example, Yang (1999).

[3]For example, during the National People's Congress and meetings of the National Committee of the CPPCC in 2004, Premier Wen Jiabao pledged to tackle the thorny issue of defaulted construction costs and wage arrears for migrant workers in the construction sector in his government work report. The Vice-minister of Labor and Social Security also declared "we will exert our utmost to tackle the farm workers' wage arrears issue and provide them with the same social security protection as their urban peers."

[4]Hirsch and Schumacher (1992) put forward the same argument for the labor market segregation of blacks in the United States.

[5]Residential Committees are neighborhood-level administrative units in China. Each Residential Committee consists of 400 to 1,000 households.

[6]We also run the specification with the martial status dummy included, but it remains insignificant.

References

Beller, A. H. (1985). "Changes in the Sex Composition of U.S. Occupations, 1960-1981", *Journal of Human Resources* 20, 235-250.

Blinder, Alan S. (1973). "Wage Discrimination: Reduced Form and Structural Estimates", *Journal of Human Resources* 8(4): 436-455.

Brown, Randall S., Moon, Marilyn, and Zoloth, Barbara S. (1980). "Incorporating Occupational Attainment in Studies of Male-Female Earnings Differntials", *Journal of Human Resources* 15(1): 3-28.

Drury, David and Arneburg, Marie W. (2001). *No More Forevers: the Chinese Labour Force in a Time of Reform*, Fato Institute for Applied Social Science.

Fan, C. Cindy (2006). "Migration and labor-market returns in urban China: results from a recent survey in Guangzhou", *Environment and Planning A* 33.

Fleisher, Belton, and Yang, Dennis Tao (2006). "Problems of China's Rural Labor

Markets and Rural-Urban Migration", *The Chinese economy* 39.
Green, D.A. (1999). "Immigrant Occupational attainment: Assimilation and Mobility over Time", *Journal of Labor Economics* 17(1): 49-79.
Hirsch, B. T. and Schumacher, E.J. (1992). "Labor Earnings, Discrimination and the Racial Composition of Jobs", *Journal of Human Resources*, 27(4): 608-28.
Lin, Justin, Wang, Gewei, and Zhao, Yaohui (2004). "Regional Inequality and Labor Migration in China", *Economic Development and Cultural Change* 52(3): 587-604.
Liu, Pak-Wai, Zhang, Jumsen, and Chong, Shu-Chuen (2004). "Occupational Segregation and Wage Differentials between Natives and Immigrants: Evidence from Hong Kong", *Journal of Development Economics* 73, 395-413.
Meng, Xin, and Zhang Junsen (2001). "The Two-Tier Labor Market in Urban China", *Journal of Comparative Economics* 29, 485-504.
Oaxaca, Ronald (1973). "Male-Female Wage Differentials in Urban Labor Market", *International Economic Review* 14(3): 693-709.
Whalley, John, and Zhang, Shunming (2004). "Inequality Change in China and (*Hukou*) Labour Mobility Restrictions", *NBER working paper series* 10683.
Wang, Meiyan (2005). "Occupation Attainment and Wage Differentials in the Urban Labor Market", *Social Science of China,* 5, 36-38.
Yang, Dennis Tao (1999). "Urban-Biased Policies and Rising Income Inequality in China", *American Economic Review*.
Zhao, Yaohui (2001). "The Role of Migrant Networks in Labor Migration: The Case of China", Working paper No. E2001012, *China Center for Economic Research*.
Zhao, Zhong (2005). "Migration, Labor Market Flexibility, and Wage Determination in China: A Review", *The Developing Economics*, June.

Table 1. Percentage distributions of resident and migrant workers by occupation-industry groups

Occupation-industry	Migrants	Residents
Blue collar in manufacture industry	64.21% (0.4795)	28.77% (0.4528)
Mean log hourly wage	1.0667 (0.5040)	1.2420 (0.5580)
White collar in manufacture industry	18.37% (0.3874)	24.36% (0.4294)
Mean log hourly wage	1.0971 (0.6514)	1.4078 (0.6769)
Blue collar in service sector industry	13.56% (0.3424)	13.79% (0.3449)
Mean log hourly wage	1.4357 (0.5475)	1.6414 (0.5851)
White collar in service sector industry	3.86% (0.1928)	33.08% (0.4706)
Mean log hourly wage	1.6666 (0.8136)	1.8222 (0.6623)
Total	100%	100%

Note: numbers in the brackets are standard deviations.

Table 2. Variable meaning and descriptive statistics

Variable	Definition	Means and Standard Deviations	
		Migrants	Residents
Log wage	Log hourly wage	1.1443 (0.5770)	1.5294 (0.6712)
Male	Dummy=1 for male	0.6141 (0.4869)	0.5345 (0.4989)
Married	Dummy=1 for currently married	0.4584 (0.4984)	0.8353 (0.3710)
Schooling	Years of schooling	9.2718 (2.5921)	11.196 (3.2556)
Experience	Years of working experience	9.0690 (8.0275)	19.568 (10.318)
Cadre	Dummy=1 for party cadre	0.0517 (0.2215)	0.3312 (0.4708)
State-owned	Dummy=1 for working in a state-owned enterprise	0.2366 (0.4251)	0.6175 (0.4861)
Allocated	Dummy=1 for getting the job through state allocation	0.0357 (0.1855)	0.2078 (0.4058)
Father-cadre	Dummy=1 if the respondent's father is a party cadre	0.1195 (0.3245)	0.3530 (0.4780)
Father-business	Dummy=1 if the respondent's father owns business	0.0470 (0.2116)	0.0191 (0.1368)
Urban-to-Urban	Dummy=1 if the respondent is a migrant has a urban *hukou* in another city	0.2069 (0.4052)	—
Beijing	Dummy=1 if the sample is from Beijing	0.2081 (0.4061)	0.3457 (0.4757)
Wuxi	Dummy=1 if the sample is from Wuxi	0.2610 (0.4393)	0.4015 (0.4903)
Sample size		1682	2204

Table 3. Multinomial logit model of industry-occupation attainment

Variable	Migrants			Residents		
	White collar in M industry	Blue collar in S industry	White collar in S industry	White collar in M industry	Blue collar in S industry	White collar in S industry
Male	-0.63532 (0.000)	-0.16138 (0.366)	0.18119 (0.581)	-0.28550 (0.023)	-0.54370 (0.001)	-0.93956 (0.000)
Schooling	0.01491 (0.625)	0.24539 (0.000)	0.33519 (0.000)	0.00214 (0.935)	0.22466 (0 .000)	0.23567 (0.000)
Experience	-0.02693 (0.378)	0.09938 (0.006)	-0.07166 (0.204)	-0.04536 (0.055)	-0.03472 (0.233)	-0.03563 (0.171)
Experience^2	0.00146 (0.068)	-0.00114 (0.243)	0.00365 (0.006)	0.00105 (0.043)	0.00138 (0.029)	0.00165 (0.004)
Married	0.53357 (0.008)	0.33675 (0.139)	0.50688 (0.220)	0.19847 (0.327)	0.36391 (0.157)	0.21525 (0.340)
Cadre	0.73545 (0.113)	2.56722 (0.000)	1.87175 (0.000)	0 .81190 (0.002)	3.00903 (0.000)	3.14551 (0.000)
Stateown	-0.19271(0.242)	-0.87515 (0.002)	0.54312 (0.124)	-0.04514 (0.730)	-0.61999 (0.000)	0.33218 (0.028)
Allocate	-1.06657(0.036)	-1.02112 (0.021)	-0.57286 (0.339)	0.07501 (0.647)	-0.28896 (0.167)	-0.04888 (0.786)
Father-cadre	0.13213 (0.590)	0.26078 (0.271)	1.09607 (0.002)	0.30552 (0.035)	0.38362 (0.029)	0.53963 (0.001)
Father-business	0.26975 (0.391)	0.65291 (0.062)	0.62222 (0.342)	-0.32912 (0.480)	-0.08331 (0.881)	-0.15577 (0.761)
Beijing	1.58736 (0.000)	-0.54177 (0.037)	0 .55342 (0.123)	-0.37801 (0.036)	-0.83130 (0.000)	-1.20716 (0.000)
Wuxi	0.88666 (0.000)	-1.14450 (0.000)	-0.92398 (0.052)	-1.49212 (0 .000)	-1.02356 (0.000)	-1.81820 (0.000)
Constant	-1.89254(0.000)	-4.52627 (0.000)	-7.06480 (0.000)	0.87510 (0.020)	-2.82366 (0.000)	-2.29381 (0.000)
N	1682			2204		

Note: Figures in parentheses indicate *p*-values. M denotes for industrial sector and S for service sector.

Table 4. Industry-occupation specific wage functions with multinomial selection

Variable	Blue collar in M industry		White collar in M industry		Blue collar in S industry		White collar in S industry	
	Migrants	Residents	Migrant	Residents	Migrants	Residents	Migrants	Residents
Male	0.17517	0.33939	0.23810	0.32416	0.06399	0.09101	0.30560	0.21389
	(0.000)	(0.000)	(0.047)	(0.000)	(0.364)	(0.129)	(0.101)	(0.000)
Schooling	0.03459	0.03639	0.03432	0.05232	0.05849	0.04666	0.09631	0.02591
	(0.000)	(0.005)	(0.034)	(0.001)	(0.008)	(0.000)	(0.145)	(0.020)
Experience	0.03697	0.01844	0.01670	0.01278	0.00793	0.01099	0.02916	0.02022
	(0.000)	(0.015)	(0.159)	(0.085)	(0.645)	(0.228)	(0.300)	(0.002)
Experience^2	-0.00098	-0.00038	-0.00042	-0.00044	0.00015	-0.00024	-0.00063	-0.00063
	(0.000)	(0.057)	(0.238)	(0.004)	(0.741)	(0.238)	(0.438)	(0.000)
Cadre	0.15283	-0.15737	0.22745	0.65197	-0.03058	0.29053	0.42381	-0.10420
	(0.398)	(0.530)	(0.306)	(0.001)	(0.880)	(0.020)	(0.144)	(0.405)
Stateown	0.14467	0.09196	0.04804	0.18833	-0.00300	-0.12899	-0.21184	-0.00677
	(0.001)	(0.032)	(0.579)	(0.000)	(0.983)	(0.193)	(0.345)	(0.914)
Beijing	-0.19902	-0.25386	0.04732	-0.42993	0.17703	-0.34362	-0.40044	-0.52194
	(0.001)	(0.007)	(0.858)	(0.000)	(0.208)	(0.000)	(0.060)	(0.000)
Wuxi	-0.07946	-0.27195	-0.02017	-0.41594	0.10625	-0.63386	-0.59968	-0.64820
	(0.037)	(0.067)	(0.915)	(0.000)	(0.498)	(0.000)	(0.057)	(0.000)
λ	0.26304	0.27781	0.12178	-0.38924	-0.26926	0.11188	0.08111	-0.37632
	(0.080)	(0.206)	(0.726)	(0.124)	(0.259)	(0.598)	(0.853)	(0.028)
Constant	0.34421	0.44021	0.36841	1.23666	0.99848	1.05377	0.19488	2.08109
	(0.000)	(0.083)	(0.489)	(0.000)	(0.100)	(0.019)	(0.905)	(0.000)
R^2	0.152267	0.213834	0.089589	0.305044	0.243422	0.281208	0.381680	0.361452
N	1080	634	309	537	228	304	65	729

Figures in parentheses indicate *p*-values. M denotes for industrial sector, and S for service sector.

Table 5. Wage decompositions

Occupation-	Observed distribution		Predicted distribution		Observed difference	Explained difference	Residual difference
industry	p^u	p^r	$\hat{p}^u$	$\hat{p}^r$	$p^u - p^r$	$p^u - \hat{p}^r$	$\hat{p}^r - p^r$
group	1	2	3	4	5	6	7
B-M	0.2877	0.6421	0.2975	0.3019	-0.3544	-0.0142	-0.3402
W-M	0.2436	0.1837	0.1368	0.2851	0.0599	-0.0415	0.1014
B-S	0.1379	0.1356	0.3875	0.2124	0.0023	-0.0745	0.0768
W-S	0.3308	0.0386	0.1782	0.2006	0.2922	0.1302	0.1620
	$\ln\overline{w}^u - \ln\overline{w}^r$	$\hat{\beta}^u(\overline{X}^u - \overline{X}^r)$	$\overline{X}^r(\hat{\beta}^u - \hat{\beta}^r)$	$p^r\hat{\beta}^u(\overline{X}^u - \overline{X}^r)$	$p^r\overline{X}^r(\hat{\beta}^u - \hat{\beta}^r)$	$\ln\overline{w}^u(p^u - \hat{p}^r)$	$\ln\overline{w}^u(\hat{p}^r - p^r)$
	8	9	10	11	12	13	14
B-M	0.1753	0.1131	0.0621	0.0726	0.0399	-0.0176	-0.4225
W-M	0.3166	0.2126	0.1040	0.0391	0.0191	-0.0584	0.1428
B-S	0.2057	-0.1115	0.3172	- 0.0151	0.0430	0.1223	0.1261
W-S	0.1562	0.1972	-0.0410	0.0761	0.0016	0.2373	0.2952
Total wage differentials			WE		WU	BE	BU
0.3851			0.1042 (27.05%)		0.1004 (26.08%)	0.0389 (10.10%)	0.1416 (36.76%)

Note: B-M denotes for blue-collar in industrial sector, W-M for white-collar in industrial sector. B-S for blue-collar in service sector, and W-S for white-collar in service sector.

Table 6. Occupational segregation: blue collar vs. white collar

Total differential	WE	WU	BE	BU
0.4087	-0.0222	0.2981	0.1326	0.0002
Percentage difference	-5.44%	72.94%	32.44 %	0.06%

Table 7. Industrial segregation: industrial sector vs. service sector

Total differential	WE	WU	BE	BU
0.5044	0.5864	-0.1789	0.0143	0.0825
Percentage difference	116.27%	-35.47%	2.84%	16.36%

Appendix 1.

(Source: Figure 1.1 of Drury and Arneburg, 2001)