



Munich Personal RePEc Archive

Formal Employment, Informal Employment and Income Differentials in Urban China

Guifu, Chen and Shigeyuki, Hamori

Xiamen University, Kobe University

September 2009

Online at <https://mpra.ub.uni-muenchen.de/17585/>

MPRA Paper No. 17585, posted 30 Sep 2009 08:27 UTC

Formal Employment, Informal Employment and Income Differentials in Urban China

Guifu Chen
School of Economics, Xiamen University, Xiamen,
China.
Email: cgf552288@yahoo.com.cn.

Shigeyuki Hamori
Faculty of Economics, Kobe University, Kobe
JAPAN
Email: hamori@econ.kobe-u.ac.jp

Abstract

Oaxaca's study (1973), along with the China Health and Nutrition Survey (CHNS) questionnaire (2004 and 2006 pooling data), is used as the basis for this study in estimating the formal-informal employment hourly income differential, as well as the formal and informal male-female employment hourly income differential in urban China. The results indicate that differences in the characteristics between formal and informal employment account for a much higher percentage of the hourly income differential than do discrimination. In addition, ignoring the sample selection bias, one finds the formal male-female, the informal male-female hourly income differential and the degree of discrimination against informal women's employment will be overestimated; conversely, the degree of discrimination against formal women's employment will be underestimated.

JEL classification number: J40, J70

Key words: formal employment, informal employment, income differentials, Chinese labor market

1. Introduction

Although informal employment has existed in China since 1949, informal employment as a fully developed concept is new. Until 1978, the ideological emphasis upon state and collective ownership severely limited the scope for other types of ownership of enterprise (for example by private, self-employed individuals), and back then informal employment existed on a small scale only. In the wake of reforms in 1978, particularly in the 1990s, other types of business ownership rapidly expanded. With the intensification of state enterprise reform from the mid-1990s onwards, and subsequent mass redundancies, the Chinese Government has taken an increasing interest in other forms of ownership of enterprise as a means of addressing the issue of unemployment. On the other hand, state and collective enterprises have also made use of informal employment, in order to cut costs. From the supply side, without taking into account mass redundancies by state- and collective-owned enterprises, the effect of heavy migrations from rural areas and new entrants into the labor market every year, has been to sharply increase informal employment.

Economists estimate informal employment based on differences between household labor survey statistics and employment figures provided by enterprises (the former is larger than the latter), based on the particular Chinese statistical system. That is, informal employment refers to the part of employment which is not recorded because enterprises do not report it for various reasons (Cai and Wang, 2004; Wu and Li, 2006). Using this methodology, the informal employment is estimated to have been about 23,130,000 in 1990, which approximate number was maintained until 1996. Informal employment rapidly increased from 1997, and rose to 174,610,000 in 2005. On the other hand, the share of informal employment also rose, and exceeded 30 per cent from 1999. The share of informal employment is shown in figure 1.

In China, persons employed in the formal sector have a much higher average income than those in informal employment. As shown in Table 1, according to the 2004 and 2006 pooling data of the China Health and Nutrition Survey (CHNS)¹ questionnaire, in 2005, the average income of those in formal employment² in urban China was 8.7 yuan an hour, while the average income of those informally employed was 5.3 yuan an hour, that is, 60.9 percent of those in formal employment. On the other hand, women have lower average earnings than men, and the male-female average earnings differentials are

¹ Source: <http://www.cpc.unc.edu/projects/china>

² The definition of formal employment in this paper will be introduced in detail in section 3.

widening³. According to CHNS, in 2005, the average formal male and informal male employment income in urban China was 8.9 yuan and 5.8 yuan an hour, respectively, while the average formal female and informal female employment income was 8.3 yuan and 4.9 yuan an hour, respectively, that is, female employment yielded 93.3 and 84.5 percent of male earnings.

This paper provides answers to the following questions: what factors affect the hourly income differentials between formal employment and informal employment, between formal male employment and formal female employment, and between informal male employment and informal female employment. In particular, how much impact does labor market discrimination have on informal employment and female hourly income?

Section 2 introduces the literature; section 3 presents the data and the definitions of variables; section 4 presents the empirical results; and the final section offers concluding remarks.

2. Selected Literature Survey

Before introducing the literature, it is essential to first clarify the characteristics of the term “informal employment”. ILO (2002) characterizes informal employment as follows: Under the expanded concept, informal employment is understood to include all remunerative work—both self-employment and wage employment—that is not recognized, regulated, or protected by existing legal or regulatory frameworks as well as non-remunerative work undertaken in an income-producing enterprise. Most informal workers—including both self-employed and wage workers—are deprived of secure work, worker’s benefits, social protection, and representation or voice. The self-employed have to take care of themselves and their enterprises. Moreover, they often face a competitive disadvantage vis-à-vis larger formal firms in capital and product markets. Informal wage workers also have to take care of themselves as they receive few (if any) employer-sponsored benefits. Moreover, both groups receive little (if any) legal or social protection. As a result of these and other factors, a higher percentage of people working in the informal economy, compared to those working in the formal economy, are poor.

In recent decades, developing countries have experienced a steady and substantial increase in the share of workers characterized by informal employment status. ILO

³ Women’s Studies Institute of China (WSIC) (2006) indicated that in 1978, the average number of earning females in urban units was 83 percent of that of males; however, it fell to 81.9 percent in 2003.

(2002) indicates informal employment as a percentage of non-agricultural employment in some regions in 2000 as follows: 48 per cent in North Africa (Algeria, 43; Morocco, 45; Tunisia, 50; Egypt, 55); 72 per cent in Sub-Saharan Africa (Benin, 93; Chad, 74; Guinea, 72; Kenya, 72; South Africa, 51); 51 per cent in Latin America (Bolivia, 63; Brazil, 60; Chile, 36 Colombia, 38; Costa Rica, 44; El Salvador, 57; Guatemala, 56; Honduras, 58; Mexico, 55; Dominican Republic, 48; Venezuela, 47); 65 per cent in Asia (India, 83; Indonesia, 78; Philippines, 72; Thailand, 51; Syria, 42).

A number of explanations have been offered to account for the rise in informal employment. Portes and coauthors provide a dynamic view of the growth of informal employment in the context of import substitution and export-oriented development strategies (Portes, 1989, 1994; Portes and Schaufler, 1993). Other studies similarly describe the linkages between export-oriented and multinational firms and informal employment through, for instance, subcontracting arrangements and export-processing zones, and thus the linkages more generally between globalization and informality (Carr and Chen, 2001; de Oliveira and Roberts, 1994; Maloney, 1997). Another explanation proffered to account for the growing share of informal employment is that higher labor standards in the formal sector may lead to a higher share of informal employment (World Bank, 1995). Whereas, in contrast, Galli and Kucera (2004) find that countries with higher labor standards tend to have higher shares of formal employment and lower shares of informal employment for 14 Latin American countries in the 1990s.

Many of the empirical studies in more recent years seem to indicate wage differences between formal and informal workers. Marcouiller, Ruiz, and Woodruff (1997) applied wage regressions to calculate unexplained wage gaps between the two sectors. The results showed that significant wage premiums are associated with work in the formal sector in El Salvador and Peru, whereas, in contrast, a premium is associated with informal work in Mexico. Tansel (2000) carried out an analysis for men and women workers separately, using the 1994 Turkish Household Expenditure Survey, defining uncovered wage earners and self-employed as part of the informal sector, while covered wage earners were considered part of the formal sector. The results indicated substantial earnings differences between the formal and informal sectors for men but not for women. Also, for Mexico, Gong and van Soest (2002) found that wage differentials between the formal and informal sectors are typically small for the lesser educated and become more significant with increasing levels of education. In addition, Pratap and Quintin (2006) find that, after controlling for selection, no wage premium remains and job satisfaction is not lower in the informal sector in the Argentinean data.

Cai and Wang (2004) provide an approach to estimating total informal employment

in urban China, based on the particular statistics system. Hu and Li (2006) indicate that about 90 per cent of the observed informal employment and formal employment monthly income differential can be explained by differences in observed characteristics in urban China 2003.

This paper will clarify what factors affect the hourly income differentials between formal and informal employment, and provide a more accurate and up-to-date measure of the hourly income differentials between formal male and formal female employment, and between informal male and informal female employment in China, taking into account the possible selectivity bias.

3. Data and the Definitions of Variables

This paper uses the data derived from the CHNS (2004 and 2006) questionnaire⁴. The questionnaires for CHNS 2004 and 2006 were distributed in 9 provinces, namely, Heilongjiang, Jiangsu, Shandong, Guizhou, Guangxi, Hubei, Henan, Hunan, and Liaoning (18 cities with 216 neighborhoods, and 36 counties with 432 villages). The data in this paper, on the other hand, make use of the urban household data only.

In section 2, multiple contending definitions of formal and informal employment are used in the empirical studies literature. Because of data availability, here we identify individuals who “*work for another person or enterprise as a permanent employee*” as being engaged in formal employment⁵. On the other hand, we define informal employment as the work of those who are “*self-employed*” or who is “*a contractor with other people or enterprise*” or “*a temporary worker*” or “*paid family worker*”. That is, in line with the Chinese economy and data availability, employment is considered formal employment when the employment is permanent, rather than being determined by if the employment is in formal enterprise.

In this paper, income includes wages, bonuses and subsidies (grocery subsidy, health allowance, bath and haircut allowance, book and newspaper allowance, housing and other subsidies). In accordance with the standard practice, the following were excluded from the analysis: students, agricultural workers, members of the armed forces, the

⁴ The data for 2003 and 2005 are pooled for our analysis. We also convert the income in 2003 into the income for 2005, taking into account the price increase (or decrease) in each province.

⁵ The 2003 and 2005 CHNS data includes the questionnaire: What is your employment position in this occupation? 1. self-employed, owner-manager with employees; 2. self-employed, independent operator with no employees; 3. works for another person or enterprise as a permanent; 4. contractor with other people or enterprise; 5. temporary worker; 6. paid family worker ; 7. unpaid family worker; 8. other; 9. unknown.

disabled, retired employees who were rehired, and male household workers. Also excluded were all persons aged 15 or less (China's labor law sets the minimum employment age at 16 years) as well as respondents who provided incomplete information on income, education, household composition, or other details. After the exclusions, the sample comprised 2,526 working individuals (1,593 men and 933 women) between the ages of 16 (school-leaving age) and 55 (state retirement age for women) or 60 (state retirement age for men)—all of whom earn incomes from a main job—and 441 females aged 16–55 doing housework (non-working women). Using the above definitions of formal and informal employment, informal employment as a percentage of non-agricultural employment is 36.9 in 2005, similar to the 36.1 obtained from the *China Statistical Book 2006*, using Cai and Wang's (2004) approach⁶.

In separate survey questions, the respondents were asked to indicate the average daily working hours and the average weekly working days. The hourly income can be calculated from the annual income and working hours. The dependent variable used in the income equations is the log of hourly cash income earned from the main job. Earnings from secondary jobs and nonmonetary benefits were excluded from the analysis.

The survey includes eight categories of education, based on academic degrees. Three education level dummies (*DS1*, *DS2*, and *DS3*) were included in hourly income equations. The duration of job training is not observed in the present data; hence, the variable of experience or tenure (years employed in the present job) could not be controlled with a direct measure. Instead, age was used, entered in both linear (*AGE*) and quadratic forms (*Age2*). Additional variables include a set of dummies representing the type of minority (*FOLK*), household registration (*HUKOU*), marital status (*MARRIED*), locality (*METRO*), settlement area (*EAST*), occupation (*TECHN*, *MANAGER*, *CRAF*, and *SEVCL*), and the type of work unit (*GOVOWN* and *COLOWN*). *EAST* is set as 1 for Jiangsu and Shandong and as 0 for elsewhere, as the average annual income of on-post staff and workers in Jiangsu and Shandong is higher in 2005 (20,957 yuan and 16,614 yuan, respectively) than in the other seven provinces⁷. *HUKOU* is set as 1 for urban household registrations and as 0 for rural household registrations. However, people who belong to households registered as rural are not rural migrants, but live in the suburban villages of the city. The definitions, means, and standard deviations of these variables are presented in Table A1 and Table A2 in the Appendix.

⁶ Definition of informal employ in this paper is not completely consistent with the one by Cai and Wang (2004).

⁷ Source: China Labor Statistical Yearbook 2006

4. Empirical Results

Table 2 presents the results of the formal and informal employment hourly income functions. The effect of all education levels is positive and significant in two equations. The coefficients of *AGE* and *AGE2* are significant in the informal employment hourly income equation; however, only *AGE* is positive and significant in the other equation. The results confirm that the hourly incomes of males are higher than those of females in both formal and informal employment. There are also significant and positive effects in relation to locality (*METRO*), settlement area (*EAST*) and occupation (*TECHN* and *CRAFT*). Being a formal government employee is associated with higher hourly incomes, but the coefficient is not significant in the other equation. On the other hand, a worker in a collective enterprise is associated with less hourly income, particularly for informal employment.

However, working women may not be randomly sampled from the overall female population. Heckman (1979) proposed a two-step correction procedure to address this problem. The first step of this procedure is to specify a participation equation for women in the form of a Probit function. Using this function, a selectivity correction factor, λ , is estimated and is included in the female income equations as a regressor. This constitutes the second step and it yields coefficients that are free from the selectivity bias due to the endogenous participation decision.

The Probit analysis on female participation is presented in Table 3. The effect of all education levels is positive and significant, indicating that women with higher education levels are more likely to participate in the labor force. As illustrated in Table 3, ethnic minorities are more likely to participate in the labor market. People from households registered as rural and living in outlying suburbs of the cities are less likely to participate in the labor market than people from households registered as urban. As women become older, the probability of participation increases, albeit at a decreasing rate. As expected, the effect of marital status for married women, as compared with single women, has a negative effect on the decision to work. The share of household members younger than 7 years or older than 65 years reduces the probability of participation. The income of other family members is positive and significant. The last two variables are included in the female participation equation, but not in the female hourly income equations, in order to achieve identification.

The results of hourly income regressions for formal and informal employment by sex are presented in Table 4. The fourth and seventh columns of Table 4 present the results of formal female and the informal female employment hourly income functions with

correction for selectivity bias, respectively. The selectivity bias terms are negative and significant. This indicates that women who have high-income opportunities, given their observed characteristics, have even better opportunities outside the wage and salary sector and are hence less likely to be included in our wage samples (Reimers, 1983). Since the selectivity bias terms are significant, OLS estimates as well as an income-differential decomposition based on OLS results would be biased.

Following Oaxaca (1973) which uses the male hourly income structure as the nondiscriminatory norm, and taking into account the possible selectivity bias, the following equation can yield

$$\ln D = \ln \bar{W}_m - \ln \bar{W}_f = (\bar{X}'_m - \bar{X}'_f) \hat{\beta}_m + \bar{X}'_f (\hat{\beta}_m - \hat{\beta}_f) - \hat{c}_f \bar{\lambda}$$

The above equation decomposes the percentage difference between the geometric means of the observed hourly income rates for the two groups into three parts: the first part is due to the differences in the average characteristics of the groups, the second is due to differences in the parameters of the income function caused by labor market discrimination and other omitted factors, and the last ($\hat{c}_f \bar{\lambda}$) is due to differences in selectivity bias⁸.

Table 5 presents the observed formal-informal, formal male-female and informal male-female hourly income differential that can be attributed to the difference in characteristics and that cannot be explained by differences in observed characteristics. The latter, due to differences in the parameters of the income function, can be attributed to labor market discrimination and other omitted variables.

The results of formal-informal hourly income differential indicate that differences in the characteristics between formal and informal employment accounts for a much higher percentage (76.35 percent) of the hourly income differential than do discrimination. On the other hand, the results of formal male-female and informal male-female hourly income differential obtained without correcting the sample selection bias indicate that discrimination accounts for a much higher percentage (141.67 percent, 80.61 percent, respectively) of the hourly income differential than do differences in the characteristics between men and women. When correcting sample selection bias, formal male-female and informal male-female hourly income differential all decrease. However, the degree of discrimination against formal women's employment rises to

⁸ The third part comprises differences between male selectivity bias and female selectivity bias ($\hat{c}_m \bar{\lambda}_m - \hat{c}_f \bar{\lambda}_f$); however, the sample of males aged 16–60 doing housework (nonworking men) is only 62; thus, we assume that male selectivity bias is zero.

169.28 percent, whereas the degree of discrimination against informal women's employment descends to 72.24 percent. That is, if ignoring the sample selection bias, the formal male-female, the informal male-female hourly income differential and the degree of discrimination against informal women's employment will be overestimated; conversely, the degree of discrimination against formal women's employment will be underestimated.

The impact of the independent variables on the explained and unexplained parts of the hourly income functions is presented in Table 6. The results of formal-informal hourly income differential indicate that education level dummies (*DS1* and *DS2*) account for a large share of the difference in characteristics between formal and informal employment, that is, approximately 0.10. On the other hand, age and the type of work unit (*GOVOWN*) are also higher at approximately 0.11 and 0.13, respectively. With respect to the unexplained part, age squared (*AGE2*) is about 0.52. This is because the coefficient of age squared in informal employment hourly income function is negative and significant, whereas it is not significant in the other equation. On the other hand, the results of analysis of formal and informal male-female employment hourly income differential indicate that age (*AGE*) accounts for a large share of the unexplained differential between male and female employment, that is, approximately 1.39 and 2.92, respectively. This is interpreted in the sense that age contributes to discrimination between male and female employment in China, that is, the coefficient of age, both for formal male and informal male employment, is positive and significant, whereas that for female employment is not significant in this paper.

5. Some Concluding Remarks

Based on Oaxaca's study (1973), and using the CHNS questionnaire (2004 and 2006 pooling data), the formal-informal employment hourly income differential was estimated, and employing the Heckman two-step procedure for sample selection bias, new estimates were provided for formal and informal male-female employment hourly income differentials in urban China.

First, the results indicate that differences in the characteristics between formal and informal employment account for a much higher percentage (76.35 percent) of the hourly income differential than do discrimination.

Second, the results of formal male-female and informal male-female hourly income differential obtained without correcting the sample selection bias indicate that

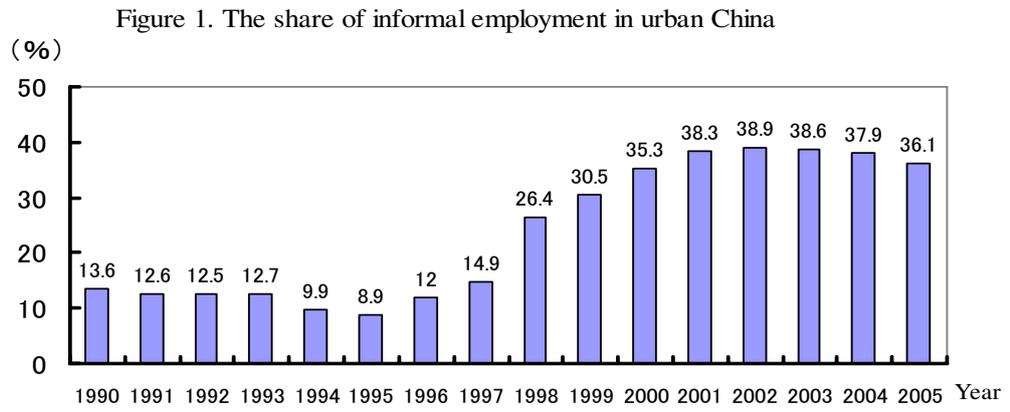
discrimination accounts for a much higher percentage (141.67 percent, 80.61 percent, respectively) of the hourly income differential than do differences in the characteristics between men and women. When correcting sample selection bias, formal male-female and informal male-female hourly income differential all decrease, however, the degree of discrimination against formal women's employment rises to 169.28 percent, whereas the degree of discrimination against informal women's employment descends to 72.24 percent.

Finally, the results of analysis of formal and informal male-female employment hourly income differential all indicate that age (*AGE*) accounts for a large share of the unexplained differential between male and female. The results indicate that the gender income differentials will decrease if the income of females rises with age through increased opportunities for promotion and the offering of other favorable treatment.

REFERENCES

- Carr, M., Chen, M. A., 2001. Globalization and the informal economy: How global trade and investment impact on the working poor. WIEGO working paper, Women in Informal Employment Globalizing and Organizing.
- Cai, F., Wang, M., 2004. The informal employment and the development of labor market. *Economic Perspectives* 2, 24–28 (in Chinese).
- de Oliveira, O., Roberts, B., 1994. The many roles of the informal sector in development: Evidence from urban labor market research, 1940–1989, in: Rakowski, C.A. (Ed.), *Contrapunto: The informal sector debate in Latin America*. Albany, NY: State University of New York Press.
- Heckman, J. J., 1979. Sample selection bias as a specification error. *Econometrica* 47, 153–161.
- Hu, W., Li, K., 2006. Informal employment and income differentials in urban China, reported in Chinese Economics Annual Conference (in Chinese).
- Galli, R., Kucera, D., 2004. Labor standards and informal employment in Latin America. *World Development* 32(5), 809–828.
- Gong, X., van Soest, 2002. Wage differentials and mobility in the urban labor market: A panel data analysis for Mexico. *Labor Economics* 9 (September), 513–529.
- ILO, 2002. Women and men in the informal economy: A statistical picture. Geneva, Switzerland: International Labor Organization.
- Maloney, W. F., 1997. Labor market structure in LDCs: Time series evidence on competing views. World Bank working paper. Washington, DC: World Bank.
- Marcouiller, D., Ruiz de Castilla, V., Woodruff, C., 1997. Formal measures of the informal-sector wage gap in Mexico, El Salvador, and Peru. *Economic Development and Cultural Change* 45(2), 367–392.
- National Bureau of Statistics of China, 2006. *China Statistical Yearbook*. China Statistics Press.
- Oaxaca, R. L., 1973. Male-female wage differentials in urban labor markets. *International Economic Review* 14, 693–709.
- Portes, A., 1989. Latin American urbanization in the years of the crisis. *Latin American Research Review*, 24(3), 7–44.
- Portes, A., 1994. When more can be less: Labor standards, development, and the informal economy, in: Rakowski, C.A. (Ed.), *Contrapunto: The informal sector debate in Latin America*. Albany, NY: State University of New York Press.
- Portes, A., Schauffler, R., 1993. Competing perspectives on the Latin American informal sector. *Population and Development Review* 19(1), 33–60.
- Pratap, S., Quintin, E., 2006. Are labor markets segmented in Argentina? A semiparametric approach. *European Economic Review* 50 (October), 1817–1841.

- Reimers, C. W., 1983. Labor market discrimination against Hispanic and black men. *The Review of Economics and Statistics* 65(4), 570–579.
- Tansel, A., 2000. Wage earners, self-employed and gender in the informal sector in Turkey. Working Paper Series no. 24 (November), Policy Research Report on Gender and Development. Washington, DC.
- Women's Studies Institute of China (WSIC), 2006. *Report on Gender Equality and Women Development in China: 1995~2005*. Social Sciences Academic Press (in Chinese).
- World Bank, 1995. Labor and economic reforms in Latin American and the Caribbean. Washington, DC: The World Bank.
- Wu, Y., Li, T., 2006. Employment situation and trends in China, in: Cai, F and Gu, B. (Ed.), *Green Book of Population and Labor*. Social Sciences Academic Press, pp. 20–43(in Chinese).



Data source: Author-compiled, based on China Statistical Book 2006.

Table 1: Employment Status, Sex, and Average Income

	Total (%)	Males (%)	females (%)	Average year income (yuan)			Average hourly income (yuan)		
				All	males	females	All	males	females
Formal employment	63.1	40.4	22.7	17 742	18235	16 864	8.7	8.9	8.3
Informal employment	36.9	18.4	18.5	11129	13236	9 027	5.3	5.8	4.9
Total	100	58.8	41.2	15 300	16668	13 345	7.4	7.9	6.8

Data source: Author-compiled, based on China Health and Nutrition Survey (CHNS) questionnaire (2004 and 2006).

TABLE 2. Results of hourly income regressions for formal employment and informal employment

Variables	Formal employment		Informal employment	
	Coefficient	<i>t</i> -value	Coefficient	<i>t</i> -value
<i>CONSTANT</i>	0.3039	1.1775	0.3025	0.9087
<i>DS1</i>	0.4222 ***	9.6990	0.3345 ***	5.2617
<i>DS2</i>	0.2692 ***	6.0851	0.1419 **	2.0830
<i>DS3</i>	0.1179 ***	2.8133	0.1885 ***	3.6668
<i>FOLK</i>	0.0335	0.6292	-0.1327	-1.5758
<i>HUKOU</i>	0.1305 **	2.2386	-0.0446	-0.7852
<i>AGE</i>	0.0259 *	1.8726	0.0452 **	2.4185
<i>AGE2</i>	-0.0002	-1.3004	-0.0006 **	-2.4265
<i>MARRIED</i>	-0.0068	-0.1176	-0.0239	-0.3185
<i>SEX</i>	0.1976 ***	6.8310	0.2542 ***	5.8852
<i>METRO</i>	0.2378 ***	8.2112	0.3030 ***	7.0386
<i>EAST</i>	0.1277 ***	3.2553	0.3612 ***	7.5325
<i>TECHN</i>	0.3095 ***	7.2683	0.1679 *	1.7475
<i>MANAGER</i>	0.2942 ***	6.2745	0.0395	0.3868
<i>CRAFT</i>	0.2380 ***	5.5713	0.1789 **	2.2002
<i>SEVCL</i>	-0.0642	-1.0178	-0.2108 ***	-4.4372
<i>GOVOWN</i>	0.2050 ***	4.4827	0.0421	0.7381
<i>COLOWN</i>	-0.1807 ***	-3.0753	-0.2157 ***	-3.1380
\bar{R}^2	0.3316		0.2332	
<i>N</i>	1593		933	

Note: Numbers in parentheses are *t* ratios.

***, **, and * indicate statistical significance at the 1, 5, and 10 percent levels, respectively.

TABLE 3. Results of the Probit analysis on female participation

Working women				
(n = 1040)				
Variables	Coefficient		t value	dP/dX
<i>CONSTANT</i>	-1.8163	**	-2.0689	-0.3538
<i>DS1</i>	2.2179	***	7.7313	0.4320
<i>DS2</i>	1.2268	***	7.5544	0.2390
<i>DS3</i>	0.5244	***	4.7681	0.1021
<i>FOLK</i>	-0.6969	***	-3.2754	-0.1357
<i>HUKOU</i>	0.8728	***	8.3525	0.1700
<i>AGE</i>	0.1613	***	3.4274	0.0314
<i>AGE2</i>	-0.0025	***	-4.2023	-0.0005
<i>MARRIED</i>	-0.9024	***	-3.5460	-0.1758
<i>METRO</i>	0.0536		0.5691	0.0104
<i>EAST</i>	0.5332	***	4.0972	0.1039
<i>RATE</i>	-0.9863	***	-3.2395	-0.1921
<i>OFAMINC</i>	0.0581	***	10.5632	0.0113
Log likelihood			-512.566	
<i>N</i>			1481	

Note: ***, **, and * indicate statistical significance at the 1, 5, and 10 percent levels, respectively.

TABLE 4. Results of hourly income regressions for formal employment and informal employment by sex

Variables	Formal employment			Informal employment		
	Males	Females(1)	Females(2)	Males	Females(1)	Females(2)
<i>CONSTANT</i>	0.4629 (1.4724)	0.5754 (1.1959)	0.9532 * (1.8547)	0.0759 (0.1571)	1.0403 * (1.9696)	1.5044 *** (2.6570)
<i>DS1</i>	0.3280 *** (6.2812)	0.6379 *** (8.1875)	0.5141 *** (5.2255)	0.3690 *** (3.6608)	0.2932 *** (3.5541)	0.1747 * (1.7820)
<i>DS2</i>	0.2068 *** (3.8950)	0.4185 *** (5.2357)	0.3137 *** (3.3137)	0.0540 (0.4923)	0.1380 (1.5752)	0.0357 (0.3620)
<i>DS3</i>	0.0751 (1.4979)	0.2353 *** (3.1146)	0.1912 ** (2.4404)	0.2108 *** (2.8467)	0.1412 ** (1.9850)	0.0739 (0.9583)
<i>FOLK</i>	0.0579 (0.9075)	-0.0315 (-0.3293)	-0.0093 (-0.0970)	-0.1772 (-1.4286)	-0.1153 (-1.0075)	-0.0812 (-0.7061)
<i>HUKOU</i>	0.1318 * (1.8354)	0.0933 (0.9476)	-0.0049 (-0.0453)	0.0340 (0.4342)	-0.1129 (-1.3567)	-0.2348 ** (-2.3595)
<i>AGE</i>	0.0296 * (1.7262)	0.0054 (0.2072)	-0.0059 (-0.2205)	0.0695 ** (2.5660)	0.0052 (0.1674)	-0.0129 (-0.4042)
<i>AGE2</i>	-0.0003 (-1.4332)	0.0001 (0.3550)	0.0003 (0.8141)	-0.0009 *** (-2.6518)	-0.0000 (-0.0930)	0.0002 (0.5371)
<i>MARRIED</i>	0.0639 (0.8386)	-0.1084 (-1.1989)	-0.0850 (-0.9348)	-0.0131 (-0.1123)	-0.0614 (-0.6184)	-0.0151 (-0.1495)
<i>METRO</i>	0.2107 *** (5.7205)	0.2812 *** (6.0266)	0.2805 *** (6.0281)	0.3544 *** (5.6411)	0.2612 *** (4.3170)	0.2594 *** (4.3043)
<i>EAST</i>	0.0941 * (1.9061)	0.1776 *** (2.7474)	0.1597 ** (2.4558)	0.2881 *** (4.1668)	0.4725 *** (7.1220)	0.4325 *** (6.3134)
<i>TECHN</i>	0.3096 *** (5.9660)	0.2933 *** (3.9467)	0.2911 *** (3.9269)	0.2476 * (1.6569)	0.1105 (0.8815)	0.1198 (0.9594)
<i>MANAGER</i>	0.3789 *** (6.9100)	0.1090 (1.2385)	0.1079 (1.2294)	0.2253 * (1.7127)	-0.3328 ** (-2.0030)	-0.3474 ** (-2.0982)

Continued

TABLE 4. Continued

Variables	Formal employment						Informal employment					
	Males		Females(1)		Females(2)		Males		Females(1)		Females(2)	
<i>CRAFT</i>	0.1954	***	0.2997	***	0.2978	***	0.1282		0.2499	**	0.2424	**
	(3.6673)		(4.1301)		(4.1149)		(0.9684)		(2.3616)		(2.2994)	
<i>SEVCL</i>	-0.0958		0.0028		-0.0011		-0.3077	***	-0.1097	*	-0.1127	*
	(-1.0233)		(0.0313)		(-0.0127)		(-4.1374)		(-1.7246)		(-1.7780)	
<i>GOVOWN</i>	0.1961	***	0.2273	***	0.2297	***	-0.0024		0.1254		0.1345	*
	(3.4181)		(3.0263)		(3.0661)		(-0.0297)		(1.5380)		(1.6541)	
<i>COLOWN</i>	-0.1756	**	-0.1462		-0.1455		-0.2273	**	-0.2122	**	-0.2295	**
	(-2.3232)		(-1.5583)		(-1.5549)		(-2.3234)		(-2.1868)		(-2.3674)	
λ					-0.2378	**					-0.2462	**
					(-2.0504)						(-2.2125)	
\bar{R}^2	0.2971		0.3937		0.3971		0.2005		0.2222		0.2289	
<i>N</i>	1020		573		573		466		467		467	

Note: Numbers in parentheses are *t* ratios.

Females (2) are results of hourly income regressions for females with correction for sample selection bias.

TABLE 5. Results of the decompositions (Oaxaca method)

	Income differentials between formal and informal employment	Income differentials between formal males and females employment		Income differentials between informal males and females employment	
		With correction for sample selection bias	Without correction	With correction for sample selection bias	Without correction
Total estimated differential	0.5589	0.0791 ²	0.1315 ¹	0.1963 ²	0.2811 ¹
Endowment differences	0.4267	-0.0548 ³	-0.0548 ³	0.0545 ³	0.0545 ³
Income discrimination	0.1322	0.1339 ⁴	0.1863 ⁴	0.1418 ⁴	0.2266 ⁴
Percentage due to endowments	76.35	-69.28	-41.67	27.76	19.39
Percentage due to discrimination	23.65	169.28	141.67	72.24	80.61

Note: 1. $\ln \bar{W}_m - \ln \bar{W}_f$ 2. $\ln \bar{W}_m - \ln \bar{W}_f + \hat{c}_f \bar{\lambda}$ 3. $(\bar{X}_m - \bar{X}_f)' \hat{\beta}_m$ 4. $\bar{X}'_f (\hat{\beta}_m - \hat{\beta}_f)$

TABLE 6. Sources of hourly income differentials due to characteristics and discrimination in the corrected sample

	Income differentials between formal and informal employment		Income differentials between formal male and female employment		Income differentials between informal male and female employment	
	explained	unexplained	explained	unexplained	explained	unexplained
Hourly income differential	0.5589		0.0791		0.1963	
Of which:	0.4267	0.1322	-0.0548	0.1339	0.0545	0.1418
Due to:						
<i>DS1</i>	0.0928	0.0142	-0.0325	-0.0828	-0.0307	0.0395
<i>DS2</i>	0.0138	0.0161	-0.0041	-0.0204	-0.0032	0.0029
<i>DS3</i>	-0.0082	-0.0162	0.0008	-0.0178	0.0019	0.0308
<i>FOLK</i>	-0.0003	0.1562	-0.0010	0.0632	0.0000	-0.0902
<i>HUKOU</i>	0.0142	0.1441	-0.0007	0.1279	-0.0024	0.2309
<i>AGE</i>	0.1057	-0.7136	0.0902	1.3900	0.2210	2.9237
<i>AGE2</i>	-0.0700	0.5237	-0.0780	-0.9055	-0.2303	-1.4688
<i>MARRIED</i>	-0.0004	0.0144	0.0002	0.1328	0.0002	0.0017
<i>SEX</i>	0.0278	-0.0283				
<i>METRO</i>	-0.0034	-0.0250	-0.0177	-0.0295	0.0223	0.0334
<i>EAST</i>	-0.0115	-0.0578	-0.0015	-0.0110	0.0020	-0.0352
<i>TECHN</i>	0.0667	0.0074	-0.0375	0.0064	-0.0048	0.0079
<i>MANAGER</i>	0.0278	0.0109	0.0163	0.0298	0.0058	0.0172
<i>CRAFT</i>	0.0248	0.0045	-0.0111	-0.0222	-0.0052	-0.0110
<i>SEVCL</i>	0.0188	0.0504	0.0049	-0.0079	0.0809	-0.0927
<i>GOVOWN</i>	0.1282	0.0262	0.0103	-0.0252	-0.0000	-0.0199
<i>COLOWN</i>	-0.0002	0.0034	0.0063	-0.0036	-0.0030	0.0002

APPENDIX

Table A1. Definition of variables

Variable	Definition
<i>LOGINC</i>	Natural logarithm of hourly income
<i>DS1</i>	1 for professional school (three-year college) or higher, 0 for others
<i>DS2</i>	1 for technical or vocational degree, 0 for others
<i>DS3</i>	1 for upper middle school degree, 0 for others
<i>FOLK</i>	1 for Han, 0 for others
<i>HUKOU</i>	1 for households registered as urban, 0 for households registered as rural
<i>AGE</i>	Age in years
<i>AGE2</i>	Age squared
<i>MARRIED</i>	1 for married, 0 for others
<i>SEX</i>	1 for male, 0 for female
<i>METRO</i>	1 for metropolitan, 0 for others
<i>EAST</i>	1 for Jiangsu and Shandong, 0 for others
<i>TECHN</i>	1 for technicians, 0 for others
<i>MANAGER</i>	1 for managers, 0 for others
<i>CRAFT</i>	1 for craft workers, 0 for others
<i>SEVCL</i>	1 for service workers, 0 for others
<i>GOVOWN</i>	1 for workers in government-owned enterprises or organizations, 0 for others
<i>COLOWN</i>	1 for workers in collective enterprises, 0 for others
<i>RATE</i>	The share of household members who are younger than 7 years or older than 65 years
<i>OFAMINC</i>	Income of other family members (in thousands of RMB)
λ	Inverse of Mill's ratio, predicted from Probit equation using all observations of females

Table A2. Means and standard deviations of variables

Variable	Means and standard deviations					
	Formal	Informal	Formal		Informal	
	(All)	(All)	(Males)	(Females)	(Males)	(Females)
<i>DS1</i>	0.3817 (0.4860)	0.1618 (0.3685)	0.3461 (0.4760)	0.4450 (0.4974)	0.1202 (0.3255)	0.2034 (0.4030)
<i>DS2</i>	0.1777 (0.3823)	0.1265 (0.3326)	0.1706 (0.3763)	0.1902 (0.3928)	0.0966 (0.2957)	0.1563 (0.3635)
<i>DS3</i>	0.1601 (0.3668)	0.2294 (0.4207)	0.1637 (0.3702)	0.1536 (0.3609)	0.2339 (0.4238)	0.2248 (0.4179)
<i>FOLK</i>	0.9297 (0.2558)	0.9400 (0.2377)	0.9235 (0.2659)	0.9407 (0.2365)	0.9399 (0.2379)	0.9400 (0.2377)
<i>HUKOU</i>	0.9322 (0.2515)	0.8232 (0.3818)	0.9304 (0.2546)	0.9354 (0.2460)	0.7876 (0.4095)	0.8587 (0.3487)
<i>AGE</i>	41.1450 (9.5179)	37.0718 (9.2108)	42.2422 (9.8293)	39.1920 (8.6051)	38.6631 (9.8245)	35.4839 (8.2648)
<i>AGE2</i>	1783.4451 (766.0038)	1459.0675 (694.6279)	1880.9206 (801.6919)	1609.9285 (664.0890)	1591.1481 (765.9617)	1327.2698 (587.1323)
<i>MARRIED</i>	0.8939 (0.3081)	0.8414 (0.3655)	0.8951 (0.3066)	0.8918 (0.3109)	0.8348 (0.3718)	0.8480 (0.3594)
<i>SEX</i>	0.6403 (0.4801)	0.4995 (0.5003)				
<i>METRO</i>	0.3685 (0.4826)	0.3826 (0.4863)	0.3382 (0.4733)	0.4223 (0.4944)	0.4142 (0.4931)	0.3512 (0.4779)
<i>EAST</i>	0.1576 (0.3645)	0.2476 (0.4318)	0.1520 (0.3592)	0.1675 (0.3738)	0.2511 (0.4341)	0.2441 (0.4300)
<i>TECHN</i>	0.2681 (0.4431)	0.0525 (0.2232)	0.2245 (0.4175)	0.3456 (0.4760)	0.0429 (0.2029)	0.0621 (0.2416)
<i>MANAGER</i>	0.1375 (0.3445)	0.0429 (0.2027)	0.1529 (0.3601)	0.1100 (0.3131)	0.0558 (0.2298)	0.0300 (0.1707)

Continued

Table A2. Continued

Variable	Means and standard deviations					
	Formal		Informal		Informal	
	(All)	(All)	(Male)	(Female)	(Male)	(Female)
<i>CRAFT</i>	0.1802 (0.3844)	0.0761 (0.2653)	0.1598 (0.3666)	0.2164 (0.4122)	0.0558 (0.2298)	0.0964 (0.2954)
<i>SEVCL</i>	0.0508 (0.2198)	0.3441 (0.4753)	0.0324 (0.1770)	0.0838 (0.2773)	0.2125 (0.4095)	0.4754 (0.4999)
<i>GOVOWN</i>	0.7859 (0.4103)	0.1608 (0.3675)	0.8049 (0.3965)	0.7522 (0.4321)	0.1760 (0.3812)	0.1456 (0.3531)
<i>COLOWN</i>	0.0973 (0.2965)	0.0965 (0.2954)	0.0843 (0.2780)	0.1204 (0.3257)	0.1030 (0.3043)	0.0899 (0.2864)
λ				0.2204 (0.3229)		0.3447 (0.4100)
<i>LOGINC</i>	1.9020 (0.6376)	1.3430 (0.6884)	1.9493 (0.6119)	1.8178 (0.6732)	1.4838 (0.6969)	1.2026 (0.6508)
<i>N</i>	1593	933	1020	573	466	467

Note: Quantities in parentheses are standard deviations.