Received: 9 September 2019 | Revised: 29 April 2020 | Accepted: 30 April 2020

DOI: 10.1111/rode.12680

SPECIAL ISSUE ARTICLE

WILEY

The gender wage gap among China's rural–urban migrants

Yan Wu[1] | **Janneke Pieters**[2,3] | **Nico Heerink**[2]

[1]Institute of Urban Development, School of Economics, Nanjing Audit University, Nanjing, China

[2]Development Economics Group, Wageningen University, Wageningen, The Netherlands

[3]IZA Institute of Labor Economics, Bonn, Germany

Correspondence
Janneke Pieters, Development Economics Group, Wageningen University, Wageningen, The Netherlands.
Email: janneke.pieters@wur.nl

Funding information
Yan Wu acknowledges the financial support from the Jiangsu Philosophy and Social Science Research Grant for Universities, Jiangsu Provincial Department of Education (Grant Number: 2019SJA0346).

Abstract
In this study, we present new empirical evidence on gender wage differences among rural–urban migrants in China. We use a data set that includes migrants residing in urban communities and those living at their workplaces—the latter were not included in the previous studies. We find that the gender wage gap among migrants is 16%–18% and does not differ between migrants living at workplaces and those living in urban communities. However, gender differences in industry sorting play a more important role for migrants living at their workplaces, whereas differences in education and experience are of importance for those living in urban communities. Overall, differences in the returns to characteristics are the main driver of the gender wage gap, especially for migrants living in urban communities.

KEYWORDS
China, gender wage gap, migration

JEL CLASSIFICATION
J31; J71; O15

1 | INTRODUCTION

China experienced historical rural to urban migrations in the past decades. Rural–urban migrants increased from approximately 25 million in 1988 to 282 million in 2016 (Meng & Zhang, 2001; NBS, 2017), now making up 35% of China's working-age population or 20% of its total population. However, the rapid increase in migration opportunities did not spread equally between men and women. Despite rising off-farm employment opportunities, recent surveys by China's National Bureau of Statistics (NBS, 2017) suggest that the female share of migrants has been around 35% since 2008.

In addition to the gender disparity in migration participation, there are pronounced gender gaps in wages among China's rural–urban migrants. These have received some attention in the literature, with a few studies exploring the magnitude and determinants of the gender wage gap. According to an early investigation by Magnani and Zhu (2012), based on nationally representative data for 2002,[1] the gender wage gap among rural–urban migrants was 30% (based on the difference in log hourly wages). Another study (Qin et al. 2016), found that male migrants earned 26% higher hourly wages than their female counterparts, using more recent data for 2010.[2] These large wage gaps, as revealed by decomposition analyses in the studies by Magnani and Zhu (2012) and Qin et al. (2016), were mainly driven by unequal returns to labor characteristics, rather than observed differences in labor characteristics. Other previous studies on the gender wage gap among China's rural–urban migrants mostly relied on data collected in one or several cities. Meng (1998), for example, analyses 1995 migrant data from Jinan, the capital city of Shandong Province. She documents a 29% gender wage gap.

While the existing studies provide primary evidence on the magnitude of the gender wage gap among China's rural–urban migrants, they typically face one or more of the following limitations. First, the data used by these studies are not representative of the migrant population. Even when previous samples were nationally representative, they excluded migrants living at workplaces (i.e., in dormitories at work sites provided by employers) because the conventional migrant surveys were conducted in urban communities, covering mainly those migrants who settled down in cities (Démurger, Gurgand, Li, & Yue, 2009). Data collected in Shenzhen, one of the major migrant-receiving cities in China, indicate that a large share—more than 40%—of migrants resided in dormitories provided by employers (Tao, Hui, Wong, & Chen, 2015). Therefore, analysis based on conventional surveys may present a distorted picture of earnings and gender inequalities in the migrant population. Second, the analysis of gender wage gaps can be highly sensitive to particular methodological choices, but previous studies did not always address this issue. Specifically, Elder, Goddeeris, and Haider (2010) emphasize that the Blinder–Oaxaca decomposition method, used by most studies, often leads to differing or even contradicting conclusions, depending on the choice of the reference wage structure. In choosing the appropriate nondiscriminatory wage structure, the current consensus is to run pooled wage regressions including the gender variable (Elder et al., 2010; Jann, 2008). Still, existing studies often exclude the gender indicator as suggested by Neumark (1988), and this methodological issue received only limited attention in empirical studies (Fortin, Lemieux, & Firpo, 2011; Gelbach, 2016; Grove, Hussey, & Jetter, 2011). Finally, few studies have assessed the driving forces of gender wage gaps among China's rural–urban migrants. Magnani and Zhu (2012) suggest that education and industrial sorting are most important in terms of observed characteristics contributing to the gender wage gap. Other studies do not distinguish the contribution of different characteristics although Meng (1998) does show that only a small portion of the gender wage gap in Jinan could be accounted for by differences in sectoral sorting.

To address these limitations, we use data provided by the Rural–Urban Migration in China (RUMiC) surveys to investigate the gender wage gap among China's rural–urban migrants in 2008. RUMiC abandoned the conventional sampling frame based on urban residential communities in view of the large shares of migrants living at workplaces. Instead, the RUMiC surveys designed a workplace-based sampling approach. As a result, migrants living in workplaces such as construction sites and factory dormitories were covered properly. As shown in Section 3, the RUMiC data indicate that, in fact, the majority of migrants live at their workplaces. In addition, we conduct rigorous Blinder–Oaxaca decomposition analyses to shed light on the driving forces of the gender wage gap. We follow

Elder et al. (2010) and Jann (2008) and derive the nondiscriminatory wage structure from a pooled wage regression including a gender dummy variable. Overall, our objective is to provide rigorous new evidence on the magnitude and determinants of the gender wage gap among China's rural–urban migrants, taking into account the large group of migrants living at their workplaces.

The main findings of this study are as follows. First, the gender wage gap among migrants is 18% for those living at their workplaces as well as for migrants residing in urban communities. This is considerably lower than previously suggested estimates. Second, while the gender wage gaps are similar for the two groups of migrants, the sources of gender wage differentials do exhibit important differences. Gender differences in industry sorting play a more important role for migrants living at their workplaces, whereas differences in education and experience are of importance for those living in urban communities. Third, differences in the returns to characteristics are the main driver of the gender wage gap, especially for migrants living in urban communities.

The remainder of the paper is organized as follows. Section 2 provides a brief overview of the literature on gender and migration. Section 3 introduces the data and presents descriptive statistics. The empirical methodology is described in Section 4, and Section 5 presents the results. Section 6 concludes the paper.

2 | LITERATURE REVIEW

Development economists have long realized that the labor market status of women is vital to development. For instance, as women's earning opportunities improve, their decision-making power, family investment in children's health, and education are likely to increase (e.g., Chang, Dong, & MacPhail, 2011; Heath & Mobarak, 2015; Jensen, 2012; Majlesi, 2016). Closing the gender wage gap and improving women's labor market opportunities are therefore a policy priority in many countries around the world.

Nevertheless, the gender wage gap has been and remains substantial in most places. Fortin (2005), for example, documents that, in the 1990s, the gender wage gap in many developed countries was approximately 20%–30%, while Nordman, Robilliard, and Roubaud (2011) report gender wage gaps of 32%–54% across seven African cities. Over the past decades, many studies have focused particularly on understanding the underlying factors accounting for gender earnings differentials. These studies consistently find that gender differences in productive attributes such as schooling and work experience can only explain part of the observed gender wage gap. In addition, occupational and sectoral segregation has been identified as an important cause of the gender wage gap in the international literature (Altonji & Blank, 1999; Borrowman & Klasen, 2020; Nordman et al., 2011). Overall, however, a large part of the gender wage gap typically remains "unexplained" after controlling for a range of personal and sometimes job characteristics (Altonji & Blank, 1999; Grove et al., 2011). That is, the gender wage gap in large part reflects the fact that women have lower "returns" to education and experience than men, sometimes even when working in the same occupation or sector.

Previous studies provide several explanations for the large unexplained component of the gender wage gap. Early studies often treated it as evidence of gender discrimination—unequal pay for equally qualified workers. However, a growing number of studies note that the traditional wage regressions include mainly human capital variables and thus have omitted factors such as skills, psychological attributes, career preference, and risk attitude, which may all contribute to the gender wage gap (see Blau & Kahn, 2017 for a review). An emerging literature has focused on the role of these kinds of labor attributes, but research remains largely constrained by the lack of appropriate measurement and the availability of data.

An important part of the research in this area has been methodological in nature, focusing on the methods to decompose the contributions of different factors to the observed gender wage gaps. The main approach in the field can be dated back to the seminal work of Oaxaca and Blinder in the early 1970s (Blinder, 1973; Oaxaca, 1973). The results of the so-called Blinder–Oaxaca (B–O) decomposition, however, is highly sensitive to the choice of nondiscriminatory (or reference) wage structure—the so-called index problem (see, e.g., Fortin et al., 2011). Consequently, many studies acknowledge the arbitrariness in choosing the reference wage structure and simply report all possible decomposition results. More recently, several studies find that the wage gap decomposition can be (better) examined with the so-called pooling strategy: regressing wages on the gender indicator together with other determinates of wages rate. The resulting estimate for the gender indicator is an attractive measure of the unexplained gap (Elder et al., 2010).

In sum, the broad literature on the gender wage gap has provided evidence for various driving forces across different countries in the world, with consistent evidence that differences in productive characteristics only account for part of the wage gap. In addition, though results are always context specific, the contribution of a particular factor can still be subject to debate even when considering the same data, as a result of differences in methodological choices.

Turning to the specific case of rural–urban migrants in China, there is little literature assessing gender inequality aspects. Ye et al. (2013) provide evidence that women have been lagging behind in migration participation in China, which created a large number of separated families—causing more than 47 million wives and 61 million children being left behind in rural China due to the out-migration of male adults (Démurger, 2015). Moreover, female migrants were also found to spend fewer years in cities and return to rural areas earlier than men (Meng, 2012). Potential explanations for these gaps in migration include the fact that off-farm jobs are often physically demanding at the early stage of industrialization, and cultural norms dictating that women are primarily responsible for household work such as caring for children and elderly.

Among the limited studies on gender wage inequality among China's migrants, Magnani and Zhu (2012) use 2002 China Household Income Project (CHIP) data and find that male migrants earn 30.2% more per hour than female migrants. Qin et al. (2016) show that male migrants earn 26% higher hourly wages than their female counterparts, based on data drawn from the 2010 National Migrant Dynamic Monitoring Survey. Although the earnings of migrants are much lower than China's urban residents, the documented gender wage gaps are not smaller compared to those of nonmigrant urban residents. Zhang, Han, Liu, and Zhao (2008) show that gender earning gaps among urban locals ranged from 19% to 27% during the period 2000–2004.

According to Magnani and Zhu (2012) and Qin et al. (2016), the so-called coefficient/discrimination effect—the unequal returns to labor characteristics—explains more than 60% of the migrant gender wage gap, whereas less than 35% of gender wage differences resulted from differences in observed labor characteristics. Among the observable factors, Magnani and Zhu (2012) suggest that education and industrial sorting are the most important characteristics explaining wage gaps. Meng (1998) finds that only a small portion of the gender wage gap in Jinan could be accounted for by differences in sectoral sorting. Moreover, Magnani and Zhu (2012) find that gender wage differentials become larger for the higher income groups—indicating that the glass ceiling effect contributes to gender wage inequalities for China's migrants—consistent with Albrecht, Björklund, and Vroman (2003), Nordman and Wolff (2009), and Bhorat and Goga (2013), who use data for Sweden, Morocco, and South Africa, respectively.

Overall, the existing literature suggests that the gender wage gaps are large among China's migrants. The documented wage gaps are partially explained by gender difference in education, experience, and industrial sorting, but a relatively large proportion of the gap cannot be explained from the observed worker or job characteristics. However, the quality of data used is an important limitation in existing studies, stemming from the fact that household surveys usually apply a residence-based sampling approach to obtain their sample frames. This method is inappropriate for migrant surveys in China, as migrants are highly mobile and often have no registered address (e.g., manufacturing workers often have dormitories inside or next to their factories; construction and restaurant workers also tend to live at construction sites and restaurants). For a better assessment of the gender wage gap among rural–urban migrants, therefore, representative data are needed, which take into account the sampling of migrants living at workplaces.

3 | DATA

Our data are drawn from the RUMiC surveys. The data for the first two waves, RUMiC 2008 and RUMiC 2009, are currently available.[3] To save space, we only report results using the data of RUMiC 2008, as the results using RUMiC 2009 data are highly consistent.[4] The data have nationally representative geographical coverage, including main migrant-receiving places throughout China (Guangzhou, Dongguan, Shenzhen, Shanghai, Nanjing, Wuxi, Hangzhou, and Ningbo in the eastern coastal region; Zhengzhou, Hefei, Luoyang, Bengbu, and Wuhan in the central region; and Chengdu and Chongqing in the western region). More importantly, the RUMiC surveys resolve the difficulty in obtaining representative sample frames and cover the large number of migrants who live at their workplaces and are excluded from conventional household surveys.

In particular, household surveys usually apply the so-called residence-based sampling approach to obtain their sample frames. However, this method is inappropriate for migrant surveys in China. First, migrants are more mobile and often have no registered address. Second, many migrants live at their workplaces. To cover both groups of migrants living at residential neighborhoods and those at workplaces, the RUMiC project developed a so-called workplace-based sampling method.

The RUMiC migrant survey was conducted in 15 cities, where the new sampling method was executed as follows. Each city is first divided into 500×500 m blocks. Several blocks (equal to around 12% of the sample size for each city) are then randomly selected. Next, a workplace census is conducted in each selected block. All workplaces (including the informal ones on the street) are interviewed with questions on the industry type, the total number of workers, and the total number of migrant workers.[5] Accordingly, a list of migrants for the block, that is, the sample frame, is created. Individuals are then randomly selected from the frames acquired in the census stage (see Akgüç, Giulietti, & Zimmermann, 2014; Kong, 2010 for more details).

To be consistent with previous studies, we exclude full-time homemakers, students, people with disabilities, retirees, and other unemployed individuals from our analysis. This results in a sample of 6,448 individuals aged 16–60 years without missing information for relevant variables, including 3,943 men (60.7%) and 2,505 women (39.3%). However, self-employed migrants do not report (or are not asked about) their living place. We only have information on place of living for wage earners, who constitute 72% of the sample (4,637 migrants). Among those, 2,866 migrants (62%) report that they live at employer-provided places such as factory dormitories, back of restaurants, and construction sites, whereas the number of migrants living in urban communities is 1,771 (38%).

3.1 | Descriptive analyses

Before turning to a descriptive analysis of gender differences in our sample, we first plot the distribution of log hourly wages for migrants living in urban communities and those living at workplaces in Figure 1. The wage distribution of the group living in urban communities lies to the right of the wage distribution of those living at workplaces. According to Kolmogorov–Smirnov test, the two distributions are not drawn from the same distribution (the corrected *p*-value is 0.001).

3.1.1 | Gender differences in wages and characteristics

Table 1 presents descriptive statistics, by gender, for all migrants in panel A (i.e., including the self-employed) and for the subsample of wage earners (panel B), migrants living at workplaces (panel C), and those living in urban communities (panel D).

For the whole sample shown in panel A, female migrants on average earned 1,403 yuan per month in the year 2007, whereas their male counterparts earned 1,735 yuan. The resulting difference, 332 yuan (20.6% of male earnings), is statistically significant at the 1% level. Working hours for both women and men are high—more than 60 h a week, with women working significantly fewer hours (1.8) than men. The yielding hourly wage rate is 6.06 yuan/h for women and 7.27 yuan/h for men—the gender wage gap in log hourly wages is 16%.

This gender wage gap is considerably lower than the 30% gap in the 2002 CHIP survey as reported in Magnani and Zhu (2012) and also lower than the gap of 26% found by Qin et al. (2016) in the 2010 NMDS data (though they report the gap as a percentage of female hourly wages—measured this way, the gap is 20% in the RUMiC migrant sample). As we report in Table 1, the difference is not because

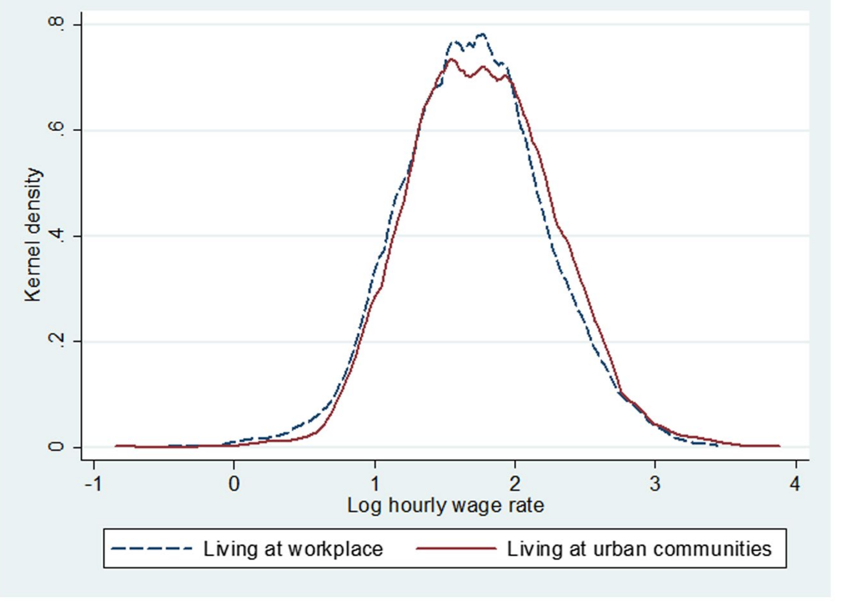

FIGURE 1 Kernel density estimates of log wage distributions

TABLE 1 Characteristics of rural–urban migrants

	Female		Male		Gender difference
	Mean	SD	Mean	SD	Mean
Panel A: whole sample					
Monthly income (yuan)	1,403	1,014	1,735	1,338	−332***
Working hours/week	61.9	17.7	63.7	17.7	−1.81***
Hourly wage (yuan)	6.06	4.28	7.27	5.49	−1.21***
Log (hourly wage) (yuan)	1.64	0.55	1.80	0.58	−0.16***
Schooling (years)	8.98	2.48	9.24	2.36	−0.26***
Age (years)	29.8	9.25	31.5	9.96	−1.70***
Off-farm experience (years)	4.25	3.69	5.02	4.64	−0.77***
Marital status (married = 1)	0.60	0.49	0.60	0.49	0.00
Number of kids: 0–16 years	0.24	0.51	0.23	0.51	0.01
Number of kids: 0–3 years	0.06	0.24	0.06	0.25	−0.00
Observations	2,505		3,943		
Panel B: wage earners					
Monthly income (yuan)	1,226	553	1,521	722	−295***
Working hours/week	57.0	14.0	59.1	14.5	−2.15***
Hourly wage (yuan)	5.70	3.21	6.85	3.85	−1.15***
Log (hourly wage) (yuan)	1.62	0.49	1.79	0.53	−0.17***
Schooling (years)	9.27	2.40	9.45	2.35	−0.18***
Age (years)	28.42	9.08	30.26	10.02	−1.84***
Off-farm experience (years)	3.57	2.93	4.51	4.36	−0.94***
Marital status (married = 1)	0.53	0.50	0.52	0.50	0.01
Number of kids: 0–16 years	0.15	0.40	0.13	0.37	0.02
Number of kids: 0–3 years	0.04	0.20	0.04	0.21	0.00
Observations	1,799		2,838		
Panel C: living at workplace					
Monthly income (yuan)	1,214	554	1,494	688	−280***
Working hours/week	59.6	14.3	60.4	14.6	−0.83
Hourly wage (yuan)	5.44	2.97	6.62	3.63	−1.18***
Log (hourly wage) (yuan)	1.57	0.48	1.76	0.52	−0.18***
Schooling (years)	9.16	2.32	9.34	2.3	−0.17
Age (years)	27.4	9.3	29.8	10.3	−2.40***
Off-farm experience (years)	3.40	2.95	4.32	4.36	−0.93***
Marital status (married = 1)	0.46	0.50	0.48	0.50	−0.02
Number of kids: 0–16 years	0.07	0.27	0.07	0.28	−0.00
Number of kids: 0–3 years	0.03	0.16	0.03	0.17	−0.00
Observations	949		1,917		
Panel D: living in urban community					
Monthly income (yuan)	1,241	552	1,576	791	−335***

(Continues)

TABLE 1 (Continued)

	Female		Male		Gender difference
	Mean	**SD**	**Mean**	**SD**	**Mean**
Working hours/week	55.56	13.1	57.61	14.18	−2.05**
Hourly wage (yuan)	6.00	3.44	7.32	4.25	−1.33***
Log (hourly wage) (yuan)	1.67	0.48	1.85	0.53	−0.18***
Schooling (years)	9.37	2.60	9.64	2.45	−0.28*
Age (years)	29.3	8.7	31.1	9.4	−1.81***
Off-farm experience (years)	3.54	2.86	4.78	4.4	−1.24***
Marital status (married = 1)	0.60	0.49	0.59	0.49	0.01
Number of kids: 0–16 years	0.24	0.48	0.23	0.48	0.00
Number of kids: 0–3 years	0.06	0.23	0.07	0.27	−0.01
Observations	850		921		

Note: *Statistically significant at the .10 level; ** at the .05 level; and *** at the .01 level.

SD, standard deviation.

Source: Calculated from RUMiC 2008.

we have a much larger fraction of migrants living at their workplaces, since among the latter we find a gender wage gap of 18%. More likely, the difference reflects the timing of the surveys (early 2000s vs. late 2000s), differences in the coverage of self-employed versus employees (in the 2002 CHIP sample, 53% of migrants is self-employed, compared to 28% in our sample).

For the sample of wage earners (panel B), both female and male migrants earn less than the migrants in the whole sample. Compared to self-employed, wage earning migrants not only work fewer hours but also have lower wage rates. Yet the gender wage gap is very similar when we consider only wage earners: the hourly wage gap equals 17% (or 20.2% of female wages).

The gender gaps in hourly wage rates for migrants living at workplaces and those living at urban communities are both about 18%, and both gaps are statistically significant (Table 1, panels C and D). For migrants living at workplaces, differences in working hours between men and women are insignificant, whereas the working hours of women for the sample of migrants living in urban communities are significantly less than men. As a result, the gender gap in monthly income is somewhat larger in the latter group.

There are also significant gender differences in human capital. Taking the whole sample, female migrants have 0.26 years less schooling, are 1.7 years younger, and have 0.77 years less off-farm experience. Among migrants living at their workplaces, women are younger and less experienced than men, but not significantly less educated. For migrants living in urban communities, gender gaps in schooling and off-farm experience are a bit larger, whereas age differences are smaller.

Comparing panels C and D of Table 1, we see that migrants living in urban communities are, on average, more educated, older, and more experienced than their counterparts living at workplaces. The former are also more likely to be married and have children. A set of t-tests on the equality of means in the core human capital variables between the two subsamples shows that their differences in education and age are significant at 1% level even though the difference is insignificant for the off-farm experience variable (p-value is 0.155).

3.1.2 | Gender and migrants' sector and occupation choices

Table 2 displays the distribution of migrants by sector and occupation among all wage earners (i.e., the subsample for which we know their place of living).[6] The industry in which migrants work most frequently is the manufacturing (24%), followed by the hotel and catering services sector (19%) and the wholesale and retail sector (16%). The three industries in total account for 60% of migrant employment. The proportion of migrants working in construction and in the social and household service sectors is 13% and 10%, respectively.

There is, not surprisingly, a major gender difference in working in the construction industry: construction accounts for 18% of male migrant employment but only 4% of female migrant employment. On the contrary, female migrant workers are more likely to work in wholesale and retail and in hotel and catering services than male migrant workers.

As regards occupations, most migrant wage earners are employed as blue-collar workers (sales, service providers, or production/transportation workers) rather than white-collar workers (managers, professionals, or clerks). The proportion of female wage earners employed as clerks and especially as

TABLE 2 Migrants' distribution by sector and occupation (wage earners)

	All wage earners		Female		Male	
	Obs.	%	Obs.	%	Obs.	%
Sector						
Manufacturing	1,135	(24.48)	431	(23.96)	704	(24.81)
Construction	596	(12.85)	77	(4.28)	519	(18.29)
Transport and communication	186	(4.01)	28	(1.56)	158	(5.57)
Wholesale and retail	732	(15.79)	399	(22.18)	333	(11.73)
Hotel and catering services	886	(19.11)	448	(24.90)	438	(15.43)
Finance and law	16	(0.35)	4	(0.22)	12	(0.42)
Real estate	185	(3.99)	40	(2.22)	145	(5.11)
Leasing and business services	45	(0.97)	27	(1.50)	18	(0.63)
Scientific research, technical service	117	(2.52)	15	(0.83)	102	(3.59)
Public facilities management	24	(0.52)	7	(0.39)	17	(0.60)
Services: social and household	473	(10.20)	195	(10.84)	278	(9.80)
Education, health, and social welfare	153	(3.30)	89	(4.95)	64	(2.26)
Entertainment	64	(1.38)	30	(1.67)	34	(1.20)
Others	25	(0.54)	9	(0.50)	16	(0.56)
Occupation						
Managers or professionals	134	(2.89)	59	(3.28)	75	(2.64)
Clerks	268	(5.78)	149	(8.28)	119	(4.19)
Sales personnel	678	(14.62)	419	(23.29)	259	(9.13)
Service provider	1,854	(39.98)	715	(39.74)	1,139	(40.13)
Production/transportation worker	1,703	(36.73)	457	(25.40)	1,246	(43.90)

Obs., number of observations.

Source: Calculated from RUMiC 2008.

sales workers is larger than that of male migrant wage earners in such jobs. Conversely, male migrants work significantly more as production/ transportation workers.

Figures 2 and 3 show the sectoral and occupational composition, respectively, by living place and gender. Across sectors (Figure 2), migrants living at workplaces are most likely to work in the manufacturing sector (26%), hotel and catering services (22%), and construction (18%). Migrants living in urban communities are most likely to work in wholesale and retail trade (28%), followed by manufacturing (18%) and hotel and catering services (15%). We also see the general pattern that men are most strongly overrepresented in the construction sector, while women dominate wholesale and retail and hotel and catering services; the gender differences do vary somewhat between migrants living at workplaces and those in urban communities. Looking at occupations (Figure 3), we find that migrants living in urban communities are much more likely to work in a sales occupation (25%) than those living at workplaces (8%), whereas the latter are more likely to work as service providers (45%) or as production/transportation workers (39%). Gender patterns are similar across the two subsamples although occupational segregation is stronger for migrants living in urban communities.

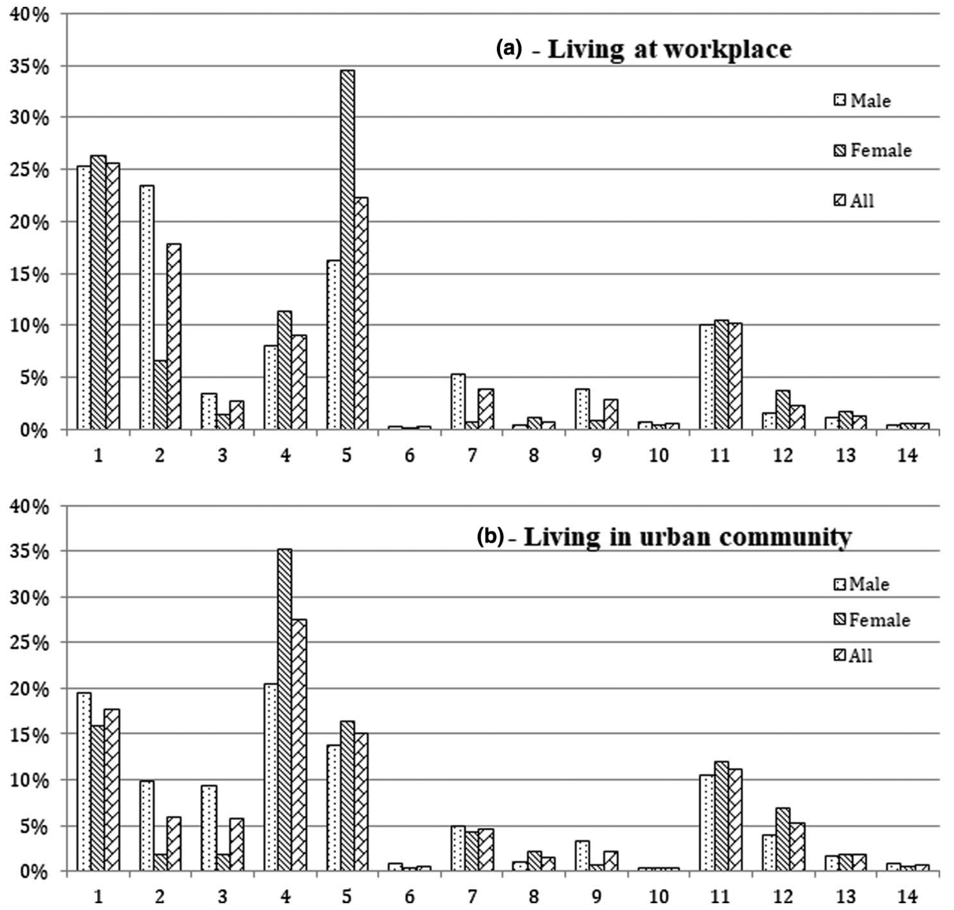

FIGURE 2 Migrants' distribution by sector. Notes: 1 Manufacturing; 2 Construction; 3 Transport and Communication; 4 Wholesale and Retail; 5 Hotel and catering Services; 6 Finance and Law; 7 Real Estate; 8 Leasing and Business Services; 9 Scientific Research and Technical Services; 10 Public Facilities Management; 11 Services: Social & Household; 12 Education Health and Social welfare; 13 Entertainment; 14 Others

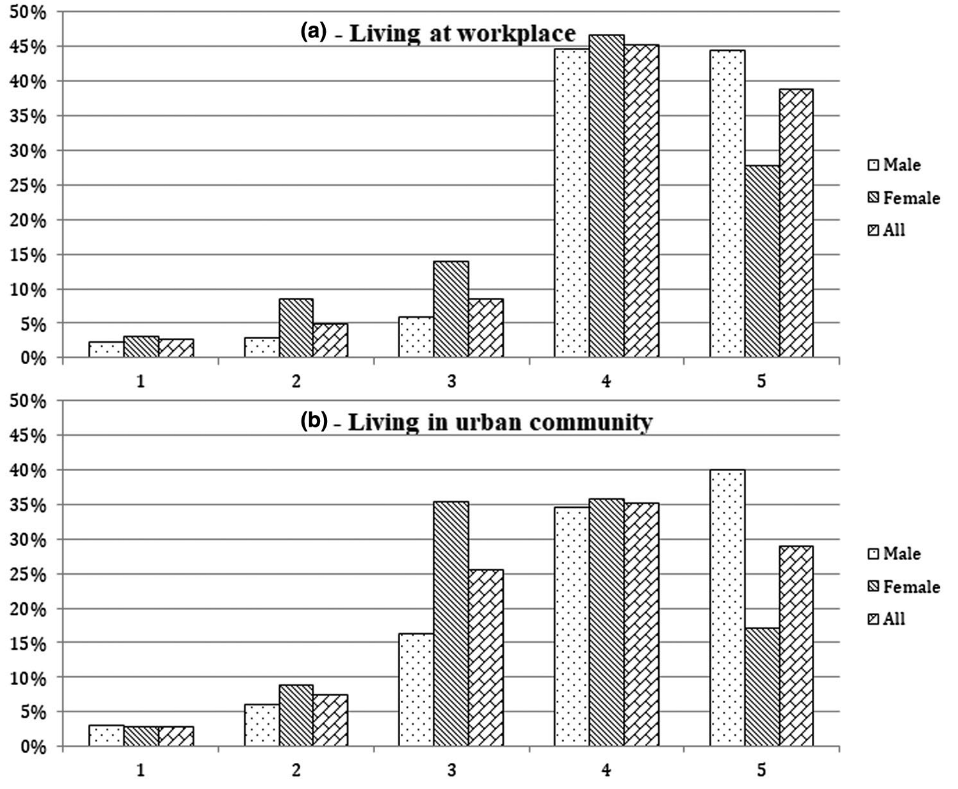

FIGURE 3 Migrants' distribution by occupation. Notes: Managers or Professionals; 2 Clerks; 3 Sales; 4 Service providers; 5 Production/Transportation Worker

3.1.3 | Gender wage differences by sector and occupation

Table 3 reports hourly wages by sector and occupation and its gender difference among the wage earners. Across all sectors, hourly wages are highest in the leasing and business services sector and the finance and law sector. Manufacturing and construction workers also receive higher wages than those in other sectors. The wage rate in hotel and catering services, which employs 25% of female migrants, is the lowest across all sectors except for the "others" sector. Female hourly wages are lower than male hourly wages within each sector. For the three major sectors of migrant employment (manufacturing, hotel and catering services, and wholesale and retail), the gender gap in wage rates is statistically significant and also economically large, about 1 yuan/h.

The evidence on wage rates by occupation suggests that the best paying occupation is that of manager or professional. Clerks earn less, but they earn more than blue-collar workers. Among the blue-collar workers, production/transportation workers earn most. Gender wage differences within all occupations are statistically significant, and they tend to be larger in the white-collar occupations.

Combing the descriptive evidence on sorting and wages, we see that the hourly wage rate is lower in sectors with a larger share of women. For instance, the hotel and catering services sector (the largest employer of female migrant workers living at their workplaces) has the second-lowest average hourly wage rate, whereas wages are relatively high in the construction sector, where male migrants are over-represented—especially in the group of migrants living at their workplaces. The descriptive evidence, albeit not definitive, suggests that the sorting of male migrants into better-paying sectors could be

TABLE 3 Hourly wage rates by sector and occupation (wage earners)

	All		Female	Male	Gender difference
	Mean	SD	Mean	Mean	Mean
Sector					
Manufacturing	7.37	3.69	6.81	7.72	–0.90***
Construction	7.23	4.04	5.82	7.45	–1.63***
Transport and communication	7.21	4.10	6.21	7.42	–1.21*
Wholesale and retail	6.10	4.88	5.65	6.57	–0.91***
Hotel and catering services	5.53	3.36	5.03	6.03	–1.00***
Finance and law	8.68	7.92	4.14	10.20	–6.05
Real estate	6.93	3.98	5.99	7.21	–1.21*
Leasing and business services	8.73	4.26	7.86	10.14	–2.28*
Scientific research, technical service	6.84	4.64	4.50	7.34	–2.83***
Public facilities management	6.22	2.81	5.70	6.42	–0.71
Services: social and household	6.39	4.30	6.47	6.32	0.15
Education, health, and social welfare	6.05	4.26	5.49	6.85	–1.35**
Entertainment	6.56	3.99	6.11	6.95	–0.84
Others	5.09	2.28	3.90	5.76	–1.86**
Occupation					
Managers or professionals	9.87	6.26	8.66	10.82	–2.16**
Clerks	7.78	4.33	6.90	8.92	–2.02***
Sales personnel	6.11	4.91	5.64	6.71	–1.07***
Service provider	5.74	3.55	5.16	6.11	–0.95***
Production/transportation worker	7.31	3.79	6.63	7.57	–0.94***

Note: *Statistically significant at the .10 level; ** at the .05 level; and *** at the .01 level.

SD, standard deviation.

Source: Calculated from RUMiC 2008.

an important factor in explaining migrant gender wage gaps, particularly for migrants living at their workplaces. Across occupations, the relationship is less clear. Large numbers of migrants, both male and female, cluster in the lower-paying occupations such as service jobs.

4 | METHODOLOGY

In this section, we discuss the details of the wage model used to estimate the effect of various worker and job characteristics on wages, which will be at the core of the decomposition analysis. After the wage model specification, we describe the decomposition analysis method in Section 4.2.

4.1 | Wage model specification

We rely on a Mincer-type wage regression model, which has been used widely in the analysis of labor markets over the past decades (see, e.g., Lemieux, 2006b) and takes the following form:

$$\ln(w_i) = \delta + \alpha F_i + \beta X_i + \varepsilon_i, \tag{1}$$

where the subscript i denotes individuals, $\ln(w)$ is the natural logarithm of the observed hourly wages,[7] δ is a constant, and ε is an error term, capturing the unobserved errors. F is a dummy variable, which is equal to 1 if the individual is a female migrant. X is a vector of productive characteristics that determines wages and β is the vector of related coefficients.

We consider two sets of X as suggested by Altonji and Blank (1999) and Lemieux (2006b). In one set, we follow the classic Mincer-type specification and include years of schooling, age, and the quadratic term of age,[8] off-farm working experience (number of years since the first time the migrant found a job in an urban area), and city dummies.[9] In the other set, we further add dummies for the worker's industry and occupation. Intuitively, sector and occupation are important job characteristics accounting for variation in wages, and we are interested to know the extent gender wage differences can be accounted for by women and men working in different sectors and occupations (and by women and men earning different wages within the same sector or occupation). There are, however, different views on whether these variables should be included. On the one hand, industry and occupation can be viewed as the outcomes of more fundamental productive characteristics such as schooling and experience, as well as gender itself, or of employer practices. On the other hand, Albrecht et al. (2003) suggest that they may reflect unmeasured human capital and may thus help explain wage differentials. We therefore choose to estimate the model with and without sector and occupation dummies, so we can present the results separately and assess how much of the wage variation is accounted for by sector and occupation, as well as how the inclusion of these variables affects the estimated coefficients for gender, schooling, age, and city.

The parameter α, the coefficient of the female dummy, is of primary interest. It represents how much the wage rates of women differ from observably identical men. A significant negative estimate for α would suggest that women are discriminated against. An unbiased estimate of α requires the following condition to hold:

$$Cov(F_i, \varepsilon_i | X_i) = 0. \tag{2}$$

This orthogonality assumption is hard to test, and it is easy to think of endogeneity issues that will violate the assumption. One commonly acknowledged problem, which in fact applies equally to the estimate of β, is sample selection bias: women and men who choose to work may be different from the unemployed or inactive population. In principle, the Heckman two-step procedure can resolve this issue, but the application of the Heckman selection model faces two difficulties in practice. First, a lack of observations of unemployed and inactive individuals[10] reduces the feasibility and efficiency of first-stage regressions. Second, the identification of the Heckman selection model requires at least one variable that affects the probability of being employed but is not directly related to an individual's wage rate. It is hard, if not impossible, to find such variables.[11] As a result, we keep the sample selection issues aside as most studies in this field do (e.g., Magnani & Zhu, 2012; Nordman et al., 2011).

Another problem, particularly relevant for the estimate of α, is the possibility of unobserved gender differences in preferences. For example, it has been reported that women are less competitive and therefore less likely to negotiate promotion into better-paying occupations, or higher wages within the same occupation (see Bertrand, 2011 for a review of the literature on gender differences in psychological attributes). Although there is little consensus on the existence and importance of such gender differences, one should keep in mind the potential omitted variable bias when interpreting the estimate of α.

4.2 | Decomposition method

The Blinder–Oaxaca decomposition method is widely used to derive the contributions of different factors to gender wage differences. However, there are various ways of implementing the method, and different choices often yield contrasting results and conclusions. We show here in detail how the choice of reference wage structures can lead to different decomposition results.

Suppose Mincer-type wage functions, estimated separately for men and women, are of the following form:

$$\ln(w_i^g) = \delta^g X_i^g + \mu_i^g, \tag{3}$$

where i denotes individuals, $\ln(w)$ is the natural logarithm of the observed hourly wage, X is the vector of productive characteristics (including a constant term), δ is the vector of coefficients, μ is an error term, and g denotes the group indicator for women (f) and men (m), respectively. The B–O approach decomposes the difference in mean wages between men and women as follows (Blinder, 1973; Oaxaca, 1973):

$$\overline{\ln(w^m)} - \overline{\ln(w^f)} = \delta^m \left(\overline{X^m} - \overline{X^f}\right) + \left(\delta^m - \delta^f\right) \overline{X^f}. \tag{4a}$$

Therefore, the mean log wage difference consists of two parts: the wage difference accounted for by different productive characteristics (the first term on the right-hand side) and the wage gap stemming from different gender-specific returns to these characteristics (the second term on the right-hand side). Different terminologies have been used to refer to the two terms: endowment effect versus coefficient effect, quantity versus price, explained versus unexplained/discrimination effect, and so on. For simplicity, we use the terms endowment effect and coefficient effect throughout the rest of the paper.

Alternatively, the difference in mean wages may be decomposed as.

$$\overline{\ln(w^m)} - \overline{\ln(w^f)} = \delta^f \left(\overline{X^m} - \overline{X^f}\right) + \left(\delta^m - \delta^f\right) \overline{X^m}. \tag{4b}$$

Here, the endowment effect is evaluated using coefficient estimates from the female sample (rather than the male sample, as in Equation 4a). In other words, the reference wage structure in Equation 4a is the male wage structure, whereas the one in Equation 4b is the female wage structure. Conversely, the coefficient effect is evaluated using male characteristics in Equation 4b, whereas it is evaluated using female characteristics in Equation 4a.

Oaxaca and Ransom (1994) suggest that using the male wage structure and the female wage structure provides the upper and lower bounds of sources of the gender wage gaps. However, Neumark (1988) argues the wage structure under nondiscrimination environments should be derived from the pooled regression estimates:

$$\overline{\ln(w^m)} - \overline{\ln(w^f)} = \delta^* \left(\overline{X^m} - \overline{X^f}\right) + \left[(\delta^m - \delta^*)\overline{X^m} + \left(\delta^* - \delta^f\right) \overline{X^f}\right], \tag{5}$$

where δ^* is the vector of coefficient estimates obtained from a wage regression using the pooled sample of women and men (without a group dummy).

As discussed by Jann (2008) and Elder et al. (2010), however, the Neumark approach tends to overstate the contribution of productive characteristics to observed wage gaps. This is because the

TABLE 4 Pooled ordinary least square estimates of log hourly wage rate

	Whole sample			Wage earner		
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Female = 1	−0.162***	−0.153***	−0.135***	−0.168***	−0.144***	−0.118***
	(0.027)	(0.020)	(0.020)	(0.025)	(0.020)	(0.019)
Years of schooling		0.047***	0.043***		0.048***	0.042***
		(0.004)	(0.004)		(0.004)	(0.004)
Age		0.056***	0.050***		0.058***	0.053***
		(0.007)	(0.007)		(0.004)	(0.005)
Age-squared/100		−0.084***	−0.076***		−0.087***	−0.081***
		(0.010)	(0.010)		(0.006)	(0.007)
Off-farm experience		0.017***	0.015***		0.023***	0.021***
		(0.003)	(0.003)		(0.003)	(0.003)
Constant	1.804***	0.649***	0.923***	1.786***	0.566***	0.824***
	(0.052)	(0.128)	(0.137)	(0.058)	(0.089)	(0.102)
City FE	No	Yes	Yes	No	Yes	Yes
Industry FE	No	No	Yes	No	No	Yes
Occupation FE	No	No	Yes	No	No	Yes
Observations	6,448	6,448	6,448	4,637	4,637	4,637
Adj. R-squared	0.019	0.223	0.262	0.025	0.313	0.363

Note: Standard errors are in parentheses and are clustered at city levels. *Statistically significant at the .10 level; ** at the .05 level; and *** at the .01 level.

omission of the group dummy variable from the pooled regression leads to omitted variable bias in the estimated coefficients δ^*, to the extent that covariates in X are correlated with group membership (i.e., gender, in our case). They suggest including the group dummy variable in the pooled wage regression model to derive the reference wage structure estimates. Therefore, we adjust the Neumark decomposition by including the gender group dummy variable in the pooled wage regression (as in the regression results reported in Tables 4 and 5).

5 | RESULTS

5.1 | Wage regression results

Table 4 provides the regression results for the whole sample and the subsample of wage earners (estimates by living place are reported in Table 5). Columns 1–3 in Table 4 present estimation results for the whole sample, whereas columns 4–6 contain results for the wage earners. In column 1, we include only the female dummy. The coefficient estimate of the female dummy variable is −0.162, indicating that the raw gender wage gap equals about 16.2%. We then add human capital characteristics, that is, education, age, and off-farm experience variables, together with city fixed effects in column 2. The estimate for the wage gap changes to −0.153. Arguably, the standard human capital traits and intercity wage differentials account for only a very small proportion of the gender wage gap for the

TABLE 5 Pooled ordinary least square estimates of log hourly wage rate by the type of residence

	Living at workplace			Living in urban community		
	(1)	(2)	(3)	(4)	(5)	(6)
Female = 1	−0.183***	−0.144***	−0.108***	−0.180***	−0.161***	−0.155***
	(0.037)	(0.026)	(0.019)	(0.033)	(0.028)	(0.030)
Years of schooling		0.042***	0.039***		0.054***	0.043***
		(0.004)	(0.004)		(0.005)	(0.006)
Age		0.059***	0.052***		0.055***	0.050***
		(0.005)	(0.006)		(0.009)	(0.009)
Age-squared/100		−0.088***	−0.080***		−0.082***	−0.075***
		(0.007)	(0.007)		(0.013)	(0.012)
Off-farm experience		0.024***	0.020***		0.024***	0.024***
		(0.003)	(0.003)		(0.005)	(0.005)
Constant	1.756***	0.642***	0.939***	1.849***	0.557***	0.791***
	(0.058)	(0.101)	(0.119)	(0.062)	(0.154)	(0.152)
City FE	No	Yes	Yes	No	Yes	Yes
Industry FE	No	No	Yes	No	No	Yes
Occupation FE	No	No	Yes	No	No	Yes
Observations	2,866	2,866	2,866	1,771	1,771	1,771
Adj. R-squared	0.028	0.321	0.391	0.030	0.314	0.351

Note: Standard errors are in parentheses and are clustered at city levels. *Statistically significant at the .10 level; ** at the .05 level; and *** at the .01 level.

whole sample. In column 3, we further add the industry and occupation dummies. The coefficient estimate for the female dummy becomes −0.135. Therefore, the sorting of male migrant workers into better-paying sectors and occupations explains part of the wage gap. However, as explained in Section 4, gender differences in industry and occupation may be the result of discriminatory practices. It is nonetheless interesting to see that the largest part of the gender wage gap remains unaccounted for after controlling for sector and occupation.

The estimated coefficients for labor attributes are in line with the findings of previous studies. Specifically, the return to schooling for rural–urban migrants is highly significant and its value is 4.2%–4.8%, comparable to the 4.1%–4.2% return rate estimated by Magnani and Zhu (2012). The coefficients for age, age-squared, and off-farm experience all have the expected signs and are highly significant too. In addition, in columns 4–6, we replicate the analyses with wage earners. The main difference from the whole sample is that including human capital characteristics and industry and occupation has larger effects on the estimates for the gender dummy and thus accounts for a larger share of the raw gender wage gaps.

In Table 5, we examine the gender wage gaps for subsamples of migrant workers living at their workplaces and those living in urban communities, excluding the self-employed migrants. We find that the raw gender gap is about 18 log points for both groups (see columns 1 and 4 in Table 5). When we include human capital control variables and industry and occupation dummies, the wage gap decreases considerably more in the sample of migrants living at their workplace—the residual wage gap for the group living at their workplaces is 0.11 versus 0.16 for the group living in urban communities (columns 3 and 6). Roughly speaking, the control variables account for 41% of the raw gender wage gap in the subsample living at workplaces and only 14% for the subsample living in urban communities. In addition, the inclusion of industry and occupation dummies accounts for a larger share of the gender wage gap among migrants living at their workplaces compared to those living in urban communities.

Table 6 shows the results of separate regressions for female and male migrants. The first noteworthy finding is that female migrants attain higher rates of return to additional schooling years than men, and this holds in all subsamples. The gender differences in the returns to schooling are more pronounced in our study than what was documented in Magnani and Zhu (2012), where education return rates are 4% for both female and male migrants.

Second, the returns to age are higher for men, and the age profile is more pronounced for men than for women, with wages increasing more steeply until a peak level at around age 30–35 for men (slightly earlier for women) and declining more at older ages. This finding applies to all the samples.

Third, the returns to additional years of off-farm experience are the same for men and women in the full sample, but larger for women than for men among wage earners, especially for the subsample not living at the workplaces (0.32 for women vs. 0.20 for men; see columns 7 and 8).

5.2 | Decomposition results

Tables 7A and 7B show the decomposition results. For robustness concerns, both Tables 7A and 7B consist of results using two different reference wage structures—the pooled coefficients as our preferred choice and the male coefficients as a robustness check. The contributions of categorical variables have all been normalized using the method proposed by Gardeazabal and Ugidos (2004).[12]

We start with the results for the endowment effects shown in Table 7A. Focusing on the results using pooled coefficients, we find that the total contribution of endowments varies quite a bit across subsamples. In the full sample, all covariates together account for only 16.6% of the raw gender wage gap. Thus, the lower wage rates received by female migrants can mostly not be accounted for by the observed productive characteristics that women bring to the labor market as compared to men, nor to differences in city, industry, and occupation sorting. For the sample of wage earners, the endowment effects explain almost 30% of the gender wage gap. In both samples, industry dummies account for the largest part of the total endowment contribution, whereas gender differences in schooling and off-farm experience also contribute to a larger wage gap. Among wage earnings, differences in experience contribute more than differences in schooling. Age and city dummies have a small negative contribution, contributing to narrow gender wage gaps. As shown in the lower panel of Table 7A, all the results are consistent when we use the male coefficients as the alternative reference wage structure.

Most interestingly, we find large differences in endowment effects for the migrants living at their workplaces (columns 5 and 6) and those living in urban communities (columns 7 and 8). First, endowment effects are much more important for migrants living at workplaces than those living in urban communities—explaining 40% of the gender gap for migrants living at their workplaces, compared to only 13% in the sample living in urban communities. The difference is driven mainly by the contributions of industry and city dummies. While city sorting contributes positively to the gender wage gap in the sample of migrants living at workplaces, it has a sizable negative contribution (i.e., narrowing the wage gap) for migrants living in urban communities. The major contributing factor for the subsample living at workplaces is the industry fixed effects, whereas industry fixed effects contribute very little in the sample living at urban communities. Our explanation is that, as shown in Figure 2, one-third of the female migrants living at their workplaces are employed in the hotel and catering services

TABLE 6 Ordinary least square estimates of log wage rate: Gender-specific estimates

	Whole sample		Wage earner		Living at workplace		Not living at workplace	
	Female	Male	Female	Male	Female	Male	Female	Male
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Years of schooling	0.046***	0.039***	0.047***	0.038***	0.043***	0.036***	0.049***	0.039***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	−0.006	−0.009
Age	0.031***	0.061***	0.026***	0.068***	0.023**	0.063***	0.019	0.078***
	(0.006)	(0.009)	(0.006)	(0.007)	(0.009)	(0.007)	−0.011	−0.013
Age-squared/100	−0.052***	−0.091***	−0.040***	−0.101***	−0.036**	−0.095***	−0.032*	−0.111***
	(0.009)	(0.013)	(0.009)	(0.010)	(0.012)	(0.010)	−0.017	−0.018
Off-farm experience	0.015***	0.015***	0.026***	0.019***	0.023***	0.019***	0.032***	0.020***
	(0.005)	(0.002)	(0.004)	(0.003)	(0.007)	(0.003)	−0.005	−0.006
Constant	1.006***	0.807***	0.992***	0.693***	1.132***	0.837***	1.048***	0.415
	(0.115)	(0.173)	(0.098)	(0.128)	(0.153)	(0.140)	−0.188	−0.244
City FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Occupation FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,505	3,943	1,799	2,838	949	1,917	850	921
Adj. R-squared	0.291	0.238	0.361	0.358	0.399	0.381	0.347	0.367

Note: Standard errors are in parentheses and are clustered at city levels. *Statistically significant at the .10 level; ** at the .05 level; and *** at the .01 level.

TABLE 7A Blinder–Oaxaca decomposition of gender wage gaps: The endowment effects

	Whole sample		Wage earner		Living at workplace		Living in urban community	
	Contribution	% of gap	Contribution	% of gap	Contribution	% of gap	Contribution	% of gap
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Reference wage structure = using pooled coefficients								
Years of schooling	0.011***	6.8%	0.007*	4.2%	0.007	3.8%	0.012*	6.7%
Age and age - squared	−0.005*	−3.1%	−0.004	−2.4%	−0.002	−1.1%	−0.001	−0.6%
Off-farm experience	0.012***	7.4%	0.021***	12.5%	0.019***	10.3%	0.030***	16.7%
Industry	0.021***	13.0%	0.026***	15.5%	0.040***	21.7%	0.002	1.1%
Occupation	0.003	1.9%	0.007	4.2%	0.008	4.3%	0.006	3.3%
City	−0.011*	−6.8%	−0.003	−1.8%	0.010	5.4%	−0.025**	−13.9%
Total	0.027**	16.7%	0.050***	29.8%	0.075***	40.8%	0.024	13.3%
Reference wage structure = using male coefficients								
Years of schooling	0.010***	6.2%	0.007*	4.2%	0.006	3.3%	0.011*	6.1%
Age and age-squared	−0.003	−1.9%	−0.002	−1.2%	−0.000	0.0%	0.004	2.2%
Off-farm experience	0.011***	6.8%	0.019***	11.3%	0.018***	9.8%	0.024***	13.3%
Industry	0.024***	14.8%	0.032***	19.0%	0.043***	23.4%	0.011	6.1%
Occupation	−0.000	0.0%	−0.002	−1.2%	0.007	3.8%	−0.008	−4.4%
City	−0.008	−4.9%	−0.002	−1.2%	0.011	6.0%	−0.021*	−11.7%
Total	0.030**	18.5%	0.046***	27.4%	0.077***	41.8%	0.022	12.2%
Raw gap	0.162***	100%	0.168***	100%	0.184***	100%	0.180***	100%

Note: *Statistically significant at the .10 level; ** at the .05 level; and *** at the .01 level.

TABLE 7B Blinder–Oaxaca decomposition of gender wage gaps: The coefficient effect

	Whole sample		Wage earner		Living at workplace		Living in urban community	
	Contribution	% of gap	Contribution	% of gap	Contribution	% of gap	Contribution	% of gap
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Reference wage structure = using pooled coefficients								
Years of schooling	−0.060	−37.0%	−0.084	−50.0%	−0.059	−32.1%	−0.089	−49.4%
Age and age-squared	0.504***	311%	0.640***	381%	0.578***	314.1%	0.973***	541%
Off-farm experience	−0.002	−1.2%	−0.026	−15.5%	−0.013	−7.1%	−0.048*	−26.7%
Industry	−0.017	−10.5%	−0.026	−15.5%	−0.050*	−27.2%	−0.015	−8.3%
Occupation	−0.029*	−17.9%	−0.026	−15.5%	0.013	7.1%	−0.062*	−34.4%
City	−0.003	−1.9%	0.001	0.6%	−0.002	−1.1%	−0.000	0.0%
Constant	−0.274	−169%	−0.375*	−223%	−0.355	−193%	−0.625*	−347%
Total	0.135***	83.3%	0.118***	70.2%	0.109***	59.2%	0.156***	86.7%
Reference wage structure = using male coefficients								
Years of schooling	−0.059	−36.4%	−0.083	−49.4%	−0.058	−31.5%	−0.087	−48.3%
Age and age-squared	0.502***	310%	0.639***	380%	0.576***	313%	0.968***	538%
Off-farm experience	−0.002	−1.2%	−0.024	−14.3%	−0.012	−6.5%	−0.043	−23.9%
Industry	−0.021	−13.0%	−0.032	−19.0%	−0.052	−28.3%	−0.025	−13.9%
Occupation	−0.026*	−16.0%	−0.017	−10.1%	0.013	7.1%	−0.048	−26.7%
City	−0.005	−3.1%	0.000	0.0%	−0.003	−1.6%	−0.004	−2.2%
Constant	−0.274	−169%	−0.375*	−223%	−0.355	−193%	−0.625*	−347%
Total	0.132	81.5%	0.122***	72.6%	0.106***	57.6%	0.158***	87.8%
Raw gap	0.162***	100%	0.168***	100%	0.184***	100%	0.180***	100%

Note: *Statistically significant at the .10 level and *** at the .01 level.

sector—the lowest-paying sector, whereas male migrants living at their workplaces are much more likely to work in the better-paid construction sector. Among migrants living in urban communities, the hotel and catering services sector accounts for a much lower share of female employment. Rather than industry sorting, human capital differences account for most of the endowment effect for migrants in urban communities.

Turning to the coefficient effects (Tables 7B), the more striking result is that gender differences in the returns to age (and its square) is the dominant factor across all samples, accounting for more than 300% of the raw gender wage gaps. As discussed earlier, the age profile is more pronounced for men than for women, with wages increasing more steeply until age 30–35 years, whereas the mean age lies between 27 and 31 years across the different subsamples.

In contrast, coefficient differences for all other observed covariates contribute negatively to the gender wage gap although most of the contributions are statistically insignificant. In addition, the constant terms show a large negative contribution, suggesting a higher conditional mean wage for women after accounting for the observed characteristics. In this sense, gender differences in unobserved characteristics and/or differences in the returns to these characteristics tend to narrow the gender wage gap.

Taken together, the coefficient effect accounts for most of the gender wage gap across all samples and is driven by the returns to age. The total coefficient contribution is lowest in the sample of migrants living at workplaces, where it accounts for 59% of the gender wage gap. Among migrants living at workplaces, human capital endowments play a relatively limited role, whereas industry sorting is an important contributor to the wage gap. Among migrants living in urban communities, industry sorting has hardly any impact, whereas schooling and off-farm experience contribute more to the gender wage gap.

5.3 | Further discussion on industrial sorting

To further explore how industrial sorting plays different roles in the two subsamples, Table 8 shows industry wage premium estimates, that is, the industry-specific wage relative to the base category (the manufacturing industry), conditional on workers' productive characteristics, the city of residence, and occupation. The estimates confirm the pattern we saw in the descriptive analysis. For migrants living at workplaces, the strongest gender segregation appears in the construction sector, where men are overrepresented, and in hotel and catering services, where women are overrepresented (Figure 2a). The regression estimates for migrants living at workplaces, in Table 8, show a significant "wage penalty" in hotel and catering services (16% lower wages compared to manufacturing), whereas there is no significant coefficient for construction.

Among migrants living in urban communities, men are more likely to work in transport and communication (Figure 2b), which has a significant negative coefficient within that subsample. Women are much less overrepresented in hotel and catering services, but instead dominate wholesale and retail, where the wages do not differ significantly from those in manufacturing. The coefficient is negative but much smaller than the negative coefficient for hotel and catering services among migrants living at workplaces.

Therefore, female migrants living at workplaces are much more likely to work in a relatively low-wage sector than their male counterparts and also than female migrants living in urban communities. On the contrary, male migrants living at workplaces are most likely work in relatively high-wage sectors (manufacturing and construction), whereas male migrants in urban communities are somewhat less concentrated in these industries.

TABLE 8 Inter-industry wage premium estimates

	Living at workplaces		Living in urban community	
	Estimate	SE	Estimate	SE
Construction	0.047	0.037	0.003	0.058
Transport and communication	0.080	0.050	−0.136	0.058**
Wholesale and retail	−0.094	0.070	−0.043	0.050
Hotel and catering services	−0.160	0.047***	−0.040	0.057
Finance and law	−0.001	0.122	0.194	0.136
Real estate	−0.018	0.072	0.051	0.060
Leasing and business services	0.186	0.083**	0.114	0.101
Scientific research, technical service	−0.128	0.031***	0.064	0.103
Public facilities management	0.054	0.055	−0.094	0.119
Services: social and household	−0.114	0.065	0.017	0.067
Education, health, and social welfare	−0.157	0.093	−0.184	0.074**
Entertainment	0.061	0.084	−0.051	0.183
Other	−0.178	0.049***	−0.383	0.158**

Note: Reference industry = manufacturing. Other explanatory variables included are female dummy, years of schooling, age, age-squared, off-farm experience, and occupation- and city-fixed effects. Standard errors are clustered at city levels. *Statistically significant at the .10 level; ** at the .05 level; and *** at the .01 level.

SE, standard error.

6 | CONCLUSION

In this study, we use a representative sample of migrants, including those living at their workplaces, to examine the magnitude and sources of gender wage disparity among rural–urban migrants in China. Our results suggest that wage rates for female migrants in China were 16%–18% lower than their male counterparts. The gaps we measure are smaller than those reported in previous studies. Arguably, the smaller gender wage gaps we document could be reflecting the evolution of China's labor market. There is evidence that the supply of rural labor has passed the Lewis turning point, and wages for men and women have been rising since 2003 (Zhang, Yang, & Wang, 2011). It could also be the case that different sampling strategies account for the differences in gender wage gaps. Yet the inclusion of migrants living at their workplaces, which constitutes the largest group of migrants in our sample but was not included in previous studies, plays a limited role: we find that the gender wage gap among migrants living in urban communities is very close to the gender wage gap for migrants living at workplaces.

Our decomposition analysis indicates that the differences in returns to labor characteristic account for the largest share of the gender wage gaps, whereas differences in endowments explain less. While this holds for different subsamples of migrants, we do document that the source of the gender wage gap differs between migrants living at their workplaces and those living in urban communities. First, differences in characteristics account for a much larger share of the gender wage gap among migrants living at workplaces (approximately 40%) compared to those living in urban communities (<14%). Second, gender differences in industry sorting play a more important role for migrants living at their workplaces, whereas differences in education and experience are more important for those living in urban communities.

These results contribute to the literature on gender aspects of rural-urban migration in China. There is very little evidence on the extent of the gender wage inequality among migrants, and we show that the wage gap may be less pronounced than that reported by previous studies. Due to their mobility and informal status, obtaining representative samples of migrants is a challenge, but with the innovative sampling strategy in the RUMiC project, this is the only available data set that properly covers migrants living at their workplaces. While the gender wage gap is similar to that for migrants living in urban communities, the underlying driving factors appear to be different.

Our study has a few important policy implications. Although the absolute level of the gender wage gap that we found is smaller than that of the conventional findings, the gender gap inequality among migrants deserves special attention. As development economists have suggested, women's working prospects are vital to the development process; better employment opportunities for women can raise their decision-making power, which can lead to increased family investment in children's health and education (e.g., Chang et al., 2011; Heath & Mobarak, 2015; Jensen, 2012; Majlesi, 2016). To mitigate the gender wage gap, our decomposition results imply that policies to promote female migrants' education and to develop higher-paying jobs that are less physically demanding and more female-friendly are particularly instrumental, as was also recommended by Magnani and Zhu (2012). Our study shows that female migrants living at their workplaces would especially benefit from better access to higher-paying jobs, whereas those living in urban communities would gain relatively more from improved education opportunities.

ACKNOWLEDGMENTS

We thank the editor and three reviewers for useful comments. The paper has also benefited from comments by participants at the international conference on "The Economic Challenges and Opportunities of Urbanization and Migration in China," held on January 11–12, 2019, at Jinan University, Guangzhou, China. Yan Wu acknowledges the financial support from the Jiangsu Philosophy and Social Science Research Grant for Universities, Jiangsu Provincial Department of Education (Grant Number: 2019SJA0346).

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in the International Data Service Center of IZA (IDSC) at http://doi.org/10.15185/izadp.7680.1.

ORCID

Yan Wu https://orcid.org/0000-0001-6239-6192
Janneke Pieters https://orcid.org/0000-0002-4575-2295
Nico Heerink https://orcid.org/0000-0002-9911-3501

ENDNOTES

1 The 2002 China Household Income Project.

2 The 2010 National Migrant Dynamics Monitoring Survey.

3 It is worth noting that the 2008 data set contains information for the calendar year 2007, whereas the 2009 data set contains information for the year 2008.

4 Results for 2009 are available upon request from the authors.

5 A migrant is defined as an individual with rural hukou and living in a city at the time of the survey.

6 The RUMiC data distinguish 28 different sectors and 23 different types of occupation. Because the number of migrant workers in some of these sectors and in some occupations is very small, and because the characteristics of some

sectors (occupations) are very similar, we aggregate the sectors into 14 different sectors and the occupations into six broad categories of occupations.

7 The use of hourly wages allows us to compare our estimates with the previous studies closest to ours. Using weekly, monthly, or annual earnings is less common in existing studies and could embed the labor supply responses (labor-leisure trade-off) to changes in hourly wage rate (see Lemieux, 2006a).

8 Experience is also often measured as age minus schooling years minus 6 (i.e., the compulsory school entry age). We use age directly as a proxy for overall labor market experience for comparing with prior studies on migrants in China.

9 Our data allow us to further control for marital status and number of children, but their inclusion has negligible effects on the empirical results.

10 It is particularly the case when using migrant data in China: migrants are highly mobile (e.g., searching/finding jobs elsewhere or returning to home villages once unemployed), and consequently migrant surveys conducted in cities have few unemployed respondents. In the RUMiC 2008 sample, for instance, the number of unemployed adults is 76 (among them, 54 are female).

11 Lee (2012) uses age-squared exclusively in the first stage regression, yielding insignificant inverse Mills ratio.

12 In addition, we follow Gelbach (2016) to attribute the changes in the estimated coefficient for the female dummy to different covariates, yielding equivalent results to the B–O decomposition results using pooled coefficients with the female dummy. The procedures and corresponding results using the Gelbach method are available upon request from the authors.

REFERENCES

Akgüç, M., Giulietti, C., & Zimmermann, K. F. (2014). The RUMiC longitudinal survey: Fostering research on labor markets in China. *IZA Journal of Labor & Development*, *3*(1), 1–14.

Albrecht, J., Björklund, A., & Vroman, S. (2003). Is there a glass ceiling in Sweden? *Journal of Labor Economics*, *21*(1), 145–177.

Altonji, J. G., & Blank, R. M. (1999). Race and gender in the labor market (Chapter 48). In O. C. Ashenfelter, & D. Card (Eds.), *Handbook of labor economics, Vol. 3, Part C* (pp. 3143–3259). Amsterdam: Elsevier.

Bertrand, M. (2011). New perspectives on gender (Chapter 17). In D. Card, & O. Ashenfelter (Eds.), *Handbook of labor economics, Vol. 4, Part B* (pp. 1543–1590). Amsterdam: Elsevier.

Bhorat, H., & Goga, S. (2013). The gender wage gap in post-Apartheid South Africa: A re-examination. *Journal of African Economies*, *22*(5), 827–848.

Blau, F. D., & Kahn, L. M. (2017). The gender wage gap: Extent, trends, and explanations. *Journal of Economic Literature*, *55*(3), 789–865.

Blinder, A. S. (1973). Wage discrimination: Reduced form and structural estimates. *The Journal of Human Resources*, *8*(4), 436–455.

Borrowman M., Klasen S. (2020). Drivers of gendered sectoral and occupational segregation in developing Countries. *Feminist Economics*, *26*, (2), 62–94. https://doi.org/10.1080/13545701.2019.1649708

Chang, H., Dong, X.-Y., & MacPhail, F. (2011). Labor migration and time use patterns of the left-behind children and elderly in rural China. *World Development*, *39*(12), 2199–2210.

Démurger, S. (2015). Migration and families left behind. *IZA World of Labor 144*. https://doi.org/10.15185/izawol.144

Démurger, S., Gurgand, M., Li, S., & Yue, X. (2009). Migrants as second-class workers in urban China? A decomposition analysis. *Journal of Comparative Economics*, *37*(4), 610–628.

Elder, T. E., Goddeeris, J. H., & Haider, S. J. (2010). Unexplained gaps and Oaxaca–blinder decompositions. *Labour Economics*, *17*(1), 284–290.

Fortin, N. M. (2005). Gender role attitudes and the labour-market outcomes of women across OECD countries. *Oxford Review of Economic Policy*, *21*(3), 416–438.

Fortin, N. M., Lemieux, T., & Firpo, S. (2011). Decomposition methods in economics (Chapter 1). In O. Ashenfelter, & D. Card (Eds.), *Handbook of labor economics, Vol. 4, Part A* (pp. 1–102). Amsterdam: Elsevier.

Gardeazabal, J., & Ugidos, A. (2004). More on identification in detailed wage decompositions. *The Review of Economics and Statistics*, *86*, 1034–1036.

Gelbach, J. B. (2016). When do covariates matter? And which ones, and how much? *Journal of Labor Economics*, *34*(2), 509–543.

Grove, W. A., Hussey, A., & Jetter, M. (2011). The gender pay gap beyond human capital: Heterogeneity in noncognitive skills and in labor market tastes. *Journal of Human Resources*, *46*(4), 827–874. https://doi.org/10.3368/jhr.46.4.827

Heath, R., & Mobarak, A. M. (2015). Manufacturing growth and the lives of Bangladeshi women. *Journal of Development Economics*, *115*(July), 1–15.

Jann, B. (2008). The Blinder-Oaxaca decomposition for linear regression models. *Stata Journal*, *8*(4), 453–479.

Jensen, R. (2012). Do labor market opportunities affect young women's work and family decisions? Experimental evidence from India. *The Quarterly Journal of Economics*, *127*(2), 753–792.

Kong, S. T. (2010). Rural–urban migration in China: Survey design and implementation. In X. Meng, & C. Manning (Eds.), with L. Shi and T. Effendi, *The great migration: Rural–urban migration in China and Indonesia*. Cheltenham, UK: Edward Elgar.

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

Lemieux, T. (2006a). Increasing residual wage inequality: Composition effects, noisy data, or rising demand for skill? *American Economic Review*, *96*(3), 461–498.

Lemieux, T. (2006b). The "mincer equation" thirty years after schooling, experience, and earnings. In S. Grossbard (Ed.), *Jacob Mincer: A pioneer of modern labor economics* (pp. 127–145). Boston, MA: Springer.

Magnani, E., & Zhu, R. (2012). Gender wage differentials among rural–urban migrants in China. *Regional Science and Urban Economics*, *42*(5), 779–793.

Majlesi, K. (2016). Labor market opportunities and women's decision making power within households. *Journal of Development Economics*, *119*, 34–47.

Meng, X. (1998). Gender occupational segregation and its impact on the gender wage differential among rural–urban migrants: A Chinese case study. *Applied Economics*, *30*(6), 741–752.

Meng, X. (2012). Labor market outcomes and reforms in China. *Journal of Economic Perspectives*, *26*(4), 75–102.

Meng, X., & Zhang, J. (2001). The two-tier labor market in urban China: Occupational segregation and wage differentials between urban residents and rural migrants in Shanghai. *Journal of Comparative Economics*, *29*(3), 485–504.

National Bureau of Statistics (NBS) (2017). *Migrants dynamic monitoring survey report for 2017*. Beijing, China: National Bureau of Statistics. Retrieved from http://www.stats.gov.cn/tjsj/zxfb/201804/t20180427_1596389.html.

Neumark, D. (1988). Employers' discriminatory behavior and the estimation of wage discrimination. *The Journal of Human Resources*, *23*(3), 279–295.

Nordman, C. J., Robilliard, A.-S., & Roubaud, F. (2011). Gender and ethnic earnings gaps in seven West African cities. *Labour Economics*, *18*(6), S132–S145.

Nordman, C. J., & Wolff, F.-C. (2009). Is there a glass ceiling in Morocco? Evidence from matched worker-firm data. *Journal of African Economies*, *18*(4), 592–633.

Oaxaca, R. (1973). Male–female wage differentials in urban labor markets. *International Economic Review*, *14*(3), 693–709.

Oaxaca, R. L., & Ransom, M. R. (1994). On discrimination and the decomposition of wage differentials. *Journal of Econometrics*, *61*(1), 5–21.

Qin, M., Brown, J. J., Padmadas, S. S., Li, B., Qi, J., & Falkingham, J. (2016). Gender inequalities in employment and wage-earning among internal labour migrants in Chinese cities. *Demographic Research*, *34*, 175–202.

Tao, L., Hui, E. C. M., Wong, F. K. W., & Chen, T. (2015). Housing choices of migrant workers in China: Beyond the Hukou perspective. *Habitat International*, *49*, 474–483.

Ye, J., Wang, C., Wu, H., He, C., & Liu, J. (2013). Internal migration and left-behind populations in China. *Journal of Peasant Studies*, *40*(6), 1119–1146.

Zhang, J., Han, J., Liu, P.-W., & Zhao, Y. (2008). Trends in the gender earnings differential in urban China, 1988–2004. *Industrial & Labor Relations Review*, *61*(2), 224–243.

Zhang, X., Yang, J., & Wang, S. (2011). China has reached the Lewis turning point. *China Economic Review*, *22*(4), 542–554.

How to cite this article: Wu Y, Pieters J, Heerink N. The gender wage gap among China's rural–urban migrants. *Rev Dev Econ*. 2020;00:1–25. https://doi.org/10.1111/rode.12680