

Routledge
Taylor & Francis Group

Journal of Chinese Economic and Business Studies

ISSN: 1476-5284 (Print) 1476-5292 (Online) Journal homepage: www.tandfonline.com/journals/rcea20

Trade unions and the wage gap between formal and informal workers: evidence from China

Xinxin Ma

To cite this article: Xinxin Ma (14 Nov 2025): Trade unions and the wage gap between formal and informal workers: evidence from China, Journal of Chinese Economic and Business Studies, DOI: 10.1080/14765284.2025.2586447

To link to this article: https://doi.org/10.1080/14765284.2025.2586447

Published online: 14 Nov 2025.

Submit your article to this journal

Article views: 73

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=rcea20

JOURNAL OF CHINESE ECONOMIC AND BUSINESS STUDIES
https://doi.org/10.1080/14765284.2025.2586447

Routledge
Taylor & Francis Group

Check for updates

Trade unions and the wage gap between formal and informal workers: evidence from China

Xinxin Ma

Faculty of Economics, Hosei University, Tokyo, Japan

ABSTRACT

Using national longitudinal survey data spanning 2014–2020 and employing a decomposition approach, this study investigates the influence of trade unions on the wage gap between formal and informal workers in China. The results reveal a union wage premium for both formal and informal workers, with the premium being larger for informal workers. The decomposition results indicate that the disparity in union membership density widens the wage gap, whereas the difference in union wage premium narrows the wage gap. These results suggest that policies aimed at expanding union coverage may increase the wages of informal workers and influence the wage gap between formal and informal workers.

ARTICLE HISTORY
Received 4 April 2025
Accepted 21 October 2025

KEYWORDS
Trade unions; wage gap; formal worker; informal worker

1. Introduction

Sustainable development is a critical issue for developing and developed countries. The 2030 Agenda for Sustainable Development outlines 17 Sustainable Development Goals aimed at creating a more sustainable world. With the rise in migrants from rural areas in developing countries, informal workers[1] have increased in recent years (e.g. [ILO] International Labor Organization 2015; 2018). As the wage gap between formal and informal workers is closely linked to inequality issues (Sustainable Development Goal 10), which may influence socially sustainable development, empirical research on the determinants of this gap has attracted global attention (e.g. Herrera-Idárraga, López-Bazo, and Motellón 2015; M. S. Li and Gao 2023; K. S. Zhang, Liu, and Ding 2018).

Based on the dual economic structure (W. A. Lewis 1954) and rural-to-urban migration models (Todaro 1969), some studies have focused on informal employment and the wage gaps between formal and informal workers in urban areas in developing countries, including China (Arias et al. 2018; Ge, Li, and Ge 2022; Herrera-Idárraga, López-Bazo, and Motellón 2015; M. S. Li and Gao 2023; K. S. Zhang, Liu, and Ding 2018). However, empirical studies examining the impact of trade unions (hereafter, unions) on the wage gap are limited (Fichtenbaum 2006; Kim and Kim 2021). The number of informal workers has increased in developing countries (Chattaraj 2016), and unionism has spread to the informal sector in South Africa (Goldman 2003; Mosoetsa 2001) and Latin America (Murillo 2000; Pencavel 1995; Rios-Avila 2017). Thus, an empirical study on the impact

CONTACT Xinxin Ma xxma@hosei.ac.jpFaculty Faculty of Economics, Hosei University, 4342 Machita-Shi Aiharamachi, Tokyo 194-0298 Japan

© 2025 The Chinese Economic Association – UK

of unions on the wage gap between formal and informal workers in developing countries has academic and practical significance.

A trade union is a type of labor market institution in an industrial society that affects wage determination in the workplace. A union has a 'monopoly face' that generates a union wage premium for union members through collective bargaining (Freeman 1980; H. G. Lewis 1963), thereby creating a wage gap or nationwide income inequality (D. T. Card 1996; Chang and Huang 2016; DiNardo, Fortin, and Lemieux 1996; Farber et al. 2021). Numerous studies have examined union wage premiums in developed (Card 1996; D. Card, Lemieux, and Riddell 2020; Choi and Ramos 2021; Farber et al. 2021; Freeman 1980; Kerr and Wittenberg 2021; H. G. Lewis 1963; Rosenfeld and Denice 2019) and developing (Gunderson, Lee, and Wang 2016; Sun and Liu 2019; Yao and Zhong 2013) countries. Some studies have demonstrated that unions affect the wage gap between groups (e.g. based on gender) (D. Card, Lemieux, and Riddell 2020; Oberfichtner, Schnabel, and Töpfer 2020) and public and private sector wage gaps (D. Card, Lemieux, and Riddell 2020; Rosenfeld and Denice 2019).

However, empirical evidence of the association between unions and the wage gap between formal and informal workers remains limited. Only two studies have investigated this relationship in the US (Fichtenbaum 2006) and South Korea (Kim and Kim 2021), and no study has focused on China – a developing and emerging market economy with a large number of formal and informal workers. Furthermore, regarding the channels through which trade unions impact the wage gap, some studies observed differences in union membership density (Fichtenbaum 2006; Kim and Kim 2021) and the union wage premium (D. T. Card 1996; D. Card, Lemieux, and Riddell 2020; Choi and Ramos 2021; Farber et al. 2021; Kerr and Wittenberg 2021; Rosenfeld and Denice 2019), which may contribute to the formation of the wage gap. However, these factors have not been investigated simultaneously. Because policy implications differ between these two channels, it is crucial to compare the magnitude of each channel. The present study attempts to fill this research gap.

Using national longitudinal survey data spanning 2014–2020 and employing decomposition methods, this study investigates the influence of trade unions on the wage gap between formal and informal workers in China. The results reveal a union wage premium for formal and informal workers, with the premium being higher for informal workers. The decomposition analysis results indicate that the disparity in union membership density (endowment effect) widens the wage gap between formal and informal workers, whereas the difference in union wage premiums (price effect) narrows the wage gap. These results suggest that policies aimed at expanding union coverage may increase the wages of informal workers and influence the wage gap between formal and informal workers.

This study substantially contributes to the literature in three ways. First, while two studies have investigated the issue in developed countries (Fichtenbaum 2006 in the US; Kim and Kim 2021 in Korea), this study is the first to focus on the issue in the context of China – the largest developing country. As the informal sector is a critical issue in developing countries (ILO 2015, 2018; Jolevski et al. 2025), this study contributes to the literature by deepening our understanding of the relationship between trade unions and informal sector development. Second, related studies (Fichtenbaum 2006; Kim and Kim 2021) have not examined the two channels of endowment (e.g. disparity in union membership density) and price (e.g. difference in union wage

premiums) effects on the wage gap simultaneously. Disentangling these effects is important because their policy implications differ. This study employs a decomposition approach to calculate the contributions of these two effects. Thus, it provides valuable academic evidence for policymakers. Third, related studies have not adequately accounted for endogeneity problems such as individual heterogeneity and sample selection biases, which may lead to biased results. To address these concerns and address individual heterogeneity, this study employs the fixed effects (FE) method. We also use a sample selection bias-adjusted model, a model using lagged variables (LV), the random effects (RE) method, and the propensity score matching (PSM) method to conduct robustness checks. Although these methods do not completely address endogeneity issues, they support the robustness of the presented evidence.

The remainder of this paper is organized as follows. section 2 reviews the related literature and discusses two channels that explain how unions influence the wage gap between formal and informal workers. section 3 introduces the data, variable settings, and the models used. section 4 presents the descriptive statistics. section 5 presents and discusses the results of the econometric analysis. Finally, section 6 presents the conclusions of this study.

2. Institutional background of trade unions in China

The functions of trade unions in China have changed with the reform of state-owned enterprises (SOEs). During the planned economic period of 1949–1977, the All-China Federation of Trade Unions (ACFTU) was formed as a national labor union organization and managed by the government. The ACFTU established branches at the regional or industrial level, and industry-level unions were managed by the regional unions (Guo and Dai 2022). The Chinese government promulgated the first *Trade Union Law* in June 1950 and published the *Resolution on the Work of the All-China Federation of Trade Unions* on 22 December 1951. The resolution stated:

> It is necessary to make all trade union workers clearly understand the party's importance to trade unions, and each level of the ACFTU at all levels must work under the unified leadership of the committees of the Communist Party of China at the same level . . . . Under the people's democratic system, the most basic and important task of trade unions is to organize and educate workers to increase labor productivity, ensuring the completion of the national production plan and striving to overfulfill to improve production, care for the daily benefits of workers, provide services to fulfill the needs of workers, and lead workers to struggle for socialism.

The union did not have the right to collectively bargain with employers for workers' wage levels during this period because the central government managed the wage setting.

The Chinese government introduced market-oriented reforms in 1978 and enforced SOE reform in the late 1990s to promote the privatization of medium- and small-sized SOEs. It granted SOEs more autonomy in wage setting and employment while promoting private sector development. Privately owned enterprises (POEs) and foreign-invested enterprises (FIEs) have developed considerably since the 1990s (Lin et al. 2020).[2]

As market-oriented reforms progressed in China, labor disputes over wages and employment increased. To address the new labor market issues, the Chinese government

promulgated the Second *Trade Union Law of the People's Republic of China* in April 1992. This was followed by the *Labor Law of the People's Republic of China*, which was adopted at the Eighth Meeting of the Standing Committee of the Eighth National People's Congress on 5 July 1994. It was issued by Presidential Order No. 28 of the People's Republic of China on 5 July 1994, and came into effect on 1 January 1995. The *Labor Contract Law of the People's Republic of China* was adopted at the 28th Session of the Standing Committee of the Tenth National People's Congress of the People's Republic of China on 29 June 2007, and took effect on 1 January 2008. These regulations stated that '*labor unions shall take measures to promote the implementation of the Labor Contract Law and the development of harmonious and stable employment relationships*.'

The 24th meeting of the Committee of the Ninth National People's Congress passed a revision of the *Trade Union Law of the People's Republic of China* on 27 October 2001. Article 2 of the new *Trade Union Law* added the provision that '*all organizations of the ACFTU represent the interests of employees and safeguard the legitimate rights and interests of employees in accordance with the law*.' Article 6 added the provision that '*protecting the legitimate rights and interests of employees is a basic responsibility of the ACFTU*.' Both public- and private-sector enterprises were subject to these regulations (Guo and Dai 2022; You 2017).

According to these regulations, Chinese trade unions could influence wage setting and employment through collective agreements with employers in the public and private sectors during the market-oriented reform period (post-1977).

3. Literature review

3.1. Two channels of the wage gap between formal and informal workers

Based on the economic theory, two channels may contribute to the wage gap between formal and informal workers. The first channel is the disparity in endowments between formal and informal workers (endowment effect). For example, following the human capital theory (Becker 1964), when the proportion of well-educated individuals is higher among formal workers than among informal workers, a wage gap may arise between the two groups. Union coverage may also affect the wage gap. Farber et al. (2021) and Kerr and Wittenberg (2021) demonstrated that the union density of disadvantaged groups (e.g. black or less-educated workers) is greater than that of their counterparts (e.g. white or well-educated workers) in developed countries. One function of unions is to increase the wages of union members through collective bargaining. Wage gaps may arise if union densities differ between formal and informal worker groups.

The second channel is the difference in the return on each factor (price effect). This may be due to discrimination against disadvantaged groups in the workplace or different wage determination mechanisms. For example, according to the discrimination theory (Becker 1957), when discrimination against informal workers persists in the workplace, formal workers may earn higher wages than informal workers, even with the same human capital levels, creating a wage gap. Additionally, following the labor market segmentation or dual labor market hypotheses (Dickens and Lang 1985), formal workers are in the primary labor market, where the influence of labor policies on wage setting (e.g. employment equality and minimum wage policies) is greater than in the secondary market,

where informal workers participate. As the impact of labor policies on wages among informal workers is smaller than that among formal workers, the effect of unions on the wage levels of informal workers may exceed that of formal workers. The disparity in union wage premiums may create a wage gap between the two groups.

As policy implications differ between these two channels, simultaneously distinguishing and examining their magnitudes is crucial for effective policymaking. For instance, if the primary driver is the disparity in union membership density, expanding union coverage among informal workers may narrow the wage gap. Conversely, if the primary driver is the difference in union wage premiums, promoting the bargaining ability of informal workers' unions or addressing discrimination against informal workers in the workplace may contribute to a reduction of the wage gap. This study addresses the challenge of distinguishing between these two channels to provide specific evidence for policymaking.

3.2. Related empirical studies

This study is related to the research on the determinants of the income gap between the formal and informal employment sectors. For countries besides China, Shahen et al. (2020) observed that income determination mechanisms differ between the formal and informal sectors. Some studies have investigated the income gap using the Blinder-Oaxaca (B-O) decomposition method (Blinder 1973; Oaxaca 1973) and revealed that endowment and price effects contribute to the income gap (Ramoni-Perazzi and Orlandoni-Merli 2021; Shahen et al. 2020).

In China, several studies have explored the determinants of the income gap between the formal and informal sectors (Ge, Li, and Ge 2022; M. S. Li and Gao 2023; K. S. Zhang, Liu, and Ding 2018) from three perspectives. First, some studies explored whether an income gap exists between the formal and informal sectors. Most studies have provided evidence of an income gap, suggesting that the Chinese labor market is segmented based on different employment sectors (K. S. Zhang, Liu, and Ding 2018; Ge, Li, and Ge 2022; M. S. Li and Gao 2023). Conversely, Y. J. Zhang and Qin (2015) found insignificant monthly and hourly income gaps between formal and informal sectors.

Second, while several studies have examined the effect of returns on education in formal and informal sectors, the results are mixed. For instance, K. S. Zhang, Liu, and Ding (2018) demonstrated that formal employees' return on education is higher than that of informal employees. However, M. S. Li and Gao (2023) observed the opposite.

Third, some studies have investigated the wage gap between formal and informal sectors using the basic B-O decomposition method. For example, using data from the China Labor-Force Dynamic Survey of 2014, K. S. Zhang, Liu, and Ding (2018) reported that the return on education for formal employees exceeded that for informal employees and that 70% and 30% of the wage gap could be explained by endowment and price effects, respectively. However, these studies did not consider the effect of unions on the wage gap.

This study makes key contributions to the literature on union wage premiums in developing countries. For instance, several studies have examined union wage premiums in South Africa; however, their findings are inconsistent. Ntlhola, Kwenda, and Ntuli (2019) used an FE model and reported premiums of 11.5% for males and 10.6% for females. Butcher and Rouse (2001) estimated a union wage premium close to 20%, whereas

Banerjee et al. (2008) documented an increase in premiums from 14% to 23% between 1995 and 2004. By contrast, Azam and Rospabe (2007) reported an exceptionally large premium of nearly 100% for black workers in South Africa. Additionally, Rufrancos (2019) found a 3.7% premium in Mexico, and Freeman (2009) and Abraham et al. (2009) reported premiums in Brazil ranging from 5% to 7% and 15% to 28%, respectively. Bhandari (2008) identified a 6% premium for contract workers and 25% premium for permanent workers in India. Clarke, Lee, and Chi (2007) and Torm (2014) reported premiums in Vietnam are 5% and 7%, respectively. Studies in China (e.g. Gunderson, Lee, and Wang 2016; M. Li and Xu 2014; Sun and Liu 2019; Yao and Zhong 2013) have estimated a wide range of union wage premiums ranging from 4.8% (Sun and Liu 2019) to 52.0% (M. Li and Xu 2014). However, the literature does not explicitly distinguish between formal and informal workers when examining union wage premiums, nor does it directly investigate the impact of unions on the wage gap between formal and informal workers.

Our study is closely related to two studies that focus on the relationship between unions and the wage gap between formal and informal workers. Using three-digit occupation data from the Current Population Survey of 1992–2002, Fichtenbaum (2006) investigated the association between unions and the wage gap between the primary and secondary labor markets in the US. The author found that the union wage premium in the secondary labor market exceeds that in the primary labor market. This suggests that unions may increase the wage levels of disadvantaged groups (e.g. informal or less-educated workers), thereby reducing the wage gap between the disadvantaged and advantaged worker groups. Using data from the Korean Labor and Income Panel Study during 2015–2019 and a decomposition method, Kim and Kim (2021) compared the differences in the impact of unions on wage discrimination against irregular workers in unionized and nonunionized workplaces in Korea. They found that the contribution of the price effect in unionized workplaces exceeded that in nonunionized workplaces, suggesting that unions may enhance discrimination against irregular workers. However, the studies did not directly or simultaneously investigate how the disparity in union density and differences in union premiums affect the wage gap between formal and informal workers. Relevant evidence on this issue from China remains limited. Therefore, this study aims to fill these gaps in the literature.

4. Methodology

4.1. Data and variable setting

This study uses longitudinal survey data from the China Family Panel Studies (CFPS), a national longitudinal survey conducted by Peking University since 2010, including follow-up surveys in 2012, 2014, 2016, 2018, and 2020. A baseline national survey was launched in 25 provinces, municipalities, and autonomous regions, and successful interviews were conducted with 14,960 households. Within these households, 33,600 adults (aged 16 years and above) and 8,990 youths (aged 14 years or younger) were interviewed during the first wave. As there were no questionnaire items on unions in the 2010 and 2012 surveys, we use data from the 2014, 2016, 2018, and 2020 waves (hereafter, CFPS 2014–2020). These waves contain all the necessary information for our analyses, including

individual respondents' union membership, employment status, earned income, education, occupation, industry sector, workplace ownership types, and other relevant details.

As informal workers comprise informal employees in the formal sector (e.g. contemporary or part-time workers) and workers in the self-employment sector, the samples included both formal and informal employees in the formal sector and workers in the self-employment sector. The survey excluded agricultural industry sector workers in rural areas and nonworking individuals. Following the *Labor Law of the People's Republic of China*, which sets the minimum working age at 16 years and mandatory retirement age in the public sector at 60 years, the samples aged 16–59 are used.[3] Samples with missing responses, abnormal values, and missing data are excluded. The total sample comprises 10,414 individuals, including 5,242 formal and 5,172 informal workers.

There are no universally accepted criteria for distinguishing between formal and informal workers. The ILO (2015; 2018) characterized the informal sector as comprising units with a small number of workers; however, the specific numerical threshold for the formal sector remains ambiguous.[4] According to the *Company Law of the People's Republic of China*, the National Bureau of Statistics (NBS) of China defined work units with fewer than eight workers as belonging to the self-employment sector, which is typically considered part of the informal sector.

We follow the definition of the self-employment sector by the Chinese government and refer to the definition of the informal sector provided by the ILO (2015; 2018) and previous studies (Ge, Li, and Ge 2022; M. S. Li and Gao 2023; K. S. Zhang, Liu, and Ding 2018). Accordingly, this study distinguishes between formal and informal workers based on the number of workers in the workplace and their occupations and situations while signing a labor contract. We define formal and informal workers as follows. A formal worker is defined as a civil servant or an individual employed with a signed labor contract in a workplace with more than eight workers. By contrast, an informal worker is defined as a noncivil servant employed in a workplace with more than eight workers but without a signed labor contract or as a worker employed in a workplace with fewer than eight workers. For sensitivity checks, we also redefine an informal worker as an individual employed in a workplace with fewer than 5 or 10 workers and without a signed labor contract.

In the wage function, we use the logarithm of the hourly wage[5] as the dependent variable. We also use the yearly wage in the robustness checks. To account for the effect of inflation, wage levels are adjusted using the annual consumer price index (CPI) published by the NBS of China. The CPI for 2014 is used as the baseline.

The key independent variable is the union. Based on the CFPS questionnaire items on union membership 'Are you a member of a trade union?' – we construct a union membership dummy variable (1 = union member, 0 = nonunion member).

We construct a formal worker dummy variable (1 = formal worker, 0 = informal worker) and an interaction term between the union and formal worker variables as the key independent variables.

We use the following control variables in the wage functions: (i) demographic factors, including the years of schooling, years of work experience, sex, health status, Communist Party of China membership (CPC), population registration status [*hukou*], and marital status dummies; (ii) work-related factors, including five occupation categories (manager, technician, clerk, operator, and other occupations) and a public sector dummy; and (iii)

region (west, central, and east) and survey year dummies (2014, 2016, 2018, and 2020). Additionally, we construct the CPC and five industrial sectors (manufacturing and construction, traffic and communication, retail and catering, service, and other) dummy variables and use them in the sensitivity checks.[6]

Appendix Table A1 provides the definitions and descriptive statistics of the variables used.

4.2. Model

We examine the union wage premium using a wage function. The ordinary least squares (OLS) method is expressed as follows:

$$\ln W_i = a + \beta_u U_i + \beta_{Fo} Fo_i + \beta_H H_i + u_i \tag{1}$$

where subscript *i* denotes an individual. *H* denotes a set of control variables (e.g. education and years of work experience). *lnW* is the logarithm of hourly wages and the dependent variable. *U* is the union membership dummy, and *Fo* is the formal worker dummy. β indicates the coefficients of each factor. β_u is the union wage premium. *a* is a constant, and *u* is an error term. We also include an interaction term between the union (*U*) and formal workers (*Fo*) in Eq. (1) to examine the differences in union wage premiums between formal and informal workers.

We use the following four methods to address the potential endogeneity issues in the OLS method. First, regarding individual heterogeneity (e.g. unobservable personality, abilities, and preferences), u_{it} in Eq. (1) includes individual-specific *a* and time-invariant factors (v_i) and idiosyncratic errors (ε_{it}). Individual heterogeneity may arise in the estimated results if v_i is maintained. Following the literature (e.g. Beauregard et al. 2025; D. T. Card 1996; Freeman 1980), this study uses the FE or RE models (Eq. [2]) to address this issue.

$$\ln W_{it} = a + \beta_u U_{it} + \beta_H H_{it} + v_i + \varepsilon_{it} \tag{2}$$

where subscript *t* is the survey year, v_i denotes the individual-specific and time-invariant factors, and ε_{it} is an idiosyncratic error. The Hausman specification and Breusch – Pagan Lagrange multiplier tests are conducted to evaluate the appropriate model choice between the OLS, RE, and FE specifications. The results indicate that the FE model is more suitable than the OLS and RE models (Table 2). Accordingly, the FE model is employed for the main estimations.

Second, we use the sample selection bias-adjusted method (Lee 1983) to address the sample selection bias of becoming a formal (or informal) worker. Paternal or maternal education and CPC membership are used for identification. The correction terms are calculated as $\lambda_{fo} = \phi(\beta N)/\Phi(\beta N)$ for becoming a formal worker and $\lambda_{inf} = \phi(\beta N)/[1 - \Phi(\beta N)]$ for becoming an informal worker. We add these correction terms to the wage function:

$$\ln W_i = a + \beta_u U_i + \beta_H H_i + \beta_\lambda \lambda_i + u_{it} \tag{3}$$

where λ indicates the selection bias-adjusted terms.

Third, a reverse causality problem may exist. For example, union wage premiums may affect workers' participation in unions. To address this issue, we use a model with the lagged time variables (LV) of the union dummy as follows:

$$lnW_{it} = a + \beta_u U_{it-1} + \beta_H H_{it} + u_{it} \tag{4}$$

where $t-1$ is the previous survey year (e.g. union membership in 2014 and wages in 2016).[6]

Fourth, to address the issue of nonrandom sorting into formal or informal worker groups based on several observable factors, we employ the PSM method (Rosenbaum and Rubin 1983). We use the formal worker group as the treatment group. By matching and resampling the observed data, the PSM method brings them closer to randomized experimental data, thereby minimizing selectivity bias and counterfactual states in the sample composition. The average treatment effect on the treated (ATT) is calculated using the matched samples based on several matching methods.

$$ATT = E(lnW_{it}(1) - lnW_{it}(0)|Formal = 1 \tag{5}$$

where $W_{it}(1)$ and $W_{it}(0)$ represent the logarithms of the hourly wages for a formal worker and a homogenous worker who becomes an informal worker, respectively.

Subsequently, we use the B-O style decomposition approach to examine the two channels (endowment and price effects) contributing to the wage gap. The standard B-O decomposition method is expressed as follows:

$$\overline{lnW}_{fo} - \overline{lnW}_{inf} = \sum \beta_{fo}(\bar{X}_{fo} - \bar{X}_{inf}) + \sum(\beta_{fo} - \beta_{inf})\bar{X}_{inf} \tag{6},$$

$$\overline{lnW}_{fo} - \overline{lnW}_{inf} = \sum \beta_{\text{inf}}(\bar{X}_{fo} - \bar{X}_{inf}) + \sum(\beta_{fo} - \beta_{inf})\bar{X}_{fo} \tag{7},$$

where *fo* and *inf* denote formal and informal workers, respectively; $\overline{lnW}_{fo} - \overline{lnW}_{inf}$ represents the wage gap between the two groups; X represents the factors that affect the wages, including the union membership and control variables; and β is the coefficient of each factor. $\sum \beta_{fo}(\bar{X}_{fo} - \bar{X}_{inf})$ and $\sum \beta_{\text{inf}}(\bar{X}_{fo} - \bar{X}_{inf})$ express the differences in endowment effect (e.g. disparity in union membership density), whereas $\sum(\beta_{fo} - \beta_{inf})\bar{X}_{inf}$ and $\sum(\beta_{fo} - \beta_{inf})\bar{X}_{fo}$ denote the differences in price effect (e.g. difference in union wage premium).

Considering the index number problem in the standard B-O decomposition method, we also use the Oaxaca and Ransom decomposition method (Oaxaca and Ransom 1994; hereafter, O-R method) as a robustness check:

$$\overline{Lnln}_{fo} - \overline{Lnln}_{inf} = \beta^*(\bar{X}_{fo} - \bar{X}_{inf}) + (\beta^* - \beta_{inf})\bar{X}_{inf} + (\beta_{fo} - \beta^*)\bar{X}_{fo} \tag{8},$$

where β^* is the employment-neutral coefficient estimated based on the wage functions using the overall sample. $\beta^*(\bar{X}_{fo} - \bar{X}_{inf})$ indicates the endowment effect. The sum of $(\beta^* - \beta_{inf})\bar{X}_{inf}$ and $(\beta_{fo} - \beta^*)\bar{X}_{inf}$ represents the price effect.

5. Descriptive statistical results

5.1. Differences in individual characteristics

Table 1 summarizes the descriptive statistics of the demographic and work-related characteristics of formal and informal workers.

The mean values of these characteristics differ between formal and informal workers. For example, the union membership density (the proportion of union workers in each group) of formal workers (31.0%) exceeds that of informal workers (8.5%). Formal workers have more years of schooling (14) than informal workers (11). Moreover, the proportion of formal workers in the public sector (50.2%) is higher than that of informal workers (14.4%).

As the differences in human capital, labor market segmentation by occupation, and workplace ownership may influence the wage gap between the two groups, we control for these factors in the wage functions in the econometric analysis.

5.2. Differences in wage distributions

Figure 1 shows the kernel density of the logarithmic hourly wage distribution for formal and informal workers. The proportion of formal workers with high wages exceeds that of informal workers, indicating a wage gap between the two groups. For example, the mean wage of formal workers (2.79) exceeds that of informal workers (2.45).

Table 1. Differences in individual characteristics between the formal and informal workers.

	Formal	Informal	Gap=	t-test
	(a)	(b)	(a)-(b)	*p*-value
Union	0.310	0.085	0.225***	0.000
Year of schooling	14	11	3***	0.000
Year of work experience	18	24	−6***	0.000
Female	0.436	0.476	−0.040***	0.070
Urban	0.929	0.832	0.097***	0.000
Married	0.732	0.811	−0.079**	0.003
Health	0.321	0.331	−0.010**	0.020
Occupation				
Manager	0.120	0.084	0.036***	0.000
Technician	0.266	0.110	0.156***	0.000
Clerk	0.174	0.112	0.062***	0.032
Operator	0.235	0.227	0.008	0.864
Other occupation	0.206	0.467	−0.261***	0.000
Public sector	0.502	0.164	0.338***	0.000
Year				
y2014	0.307	0.324	−0.017	0.283
y2016	0.135	0.258	−0.123***	0.000
y2018	0.398	0.305	0.093***	0.000
y2020	0.160	0.113	0.047***	0.000
Region				
East	0.554	0.429	0.125***	0.016
Central	0.290	0.374	−0.084***	0.000
West	0.156	0.198	−0.042***	0.000
No. of observations	5242	5172		

Notes: ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

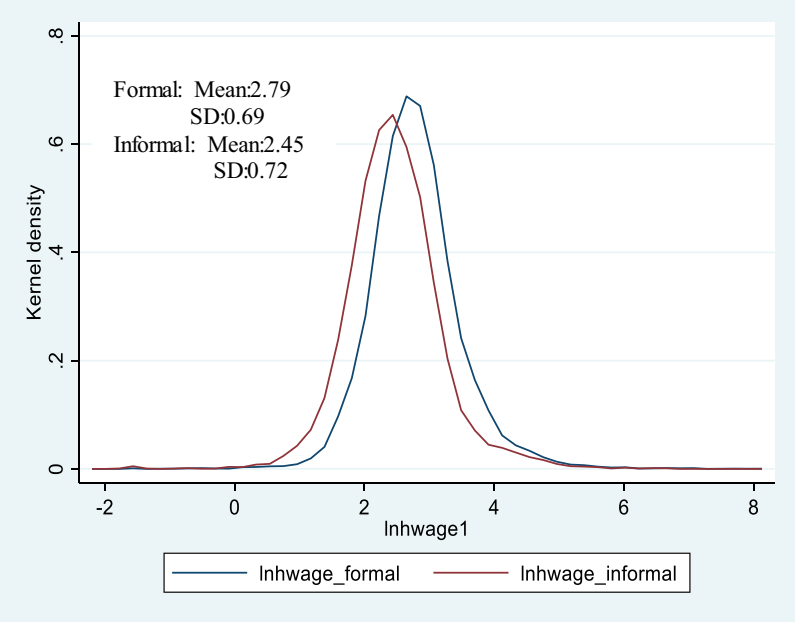

Figure 1. Kernel density of wage distributions among the formal and informal workers.

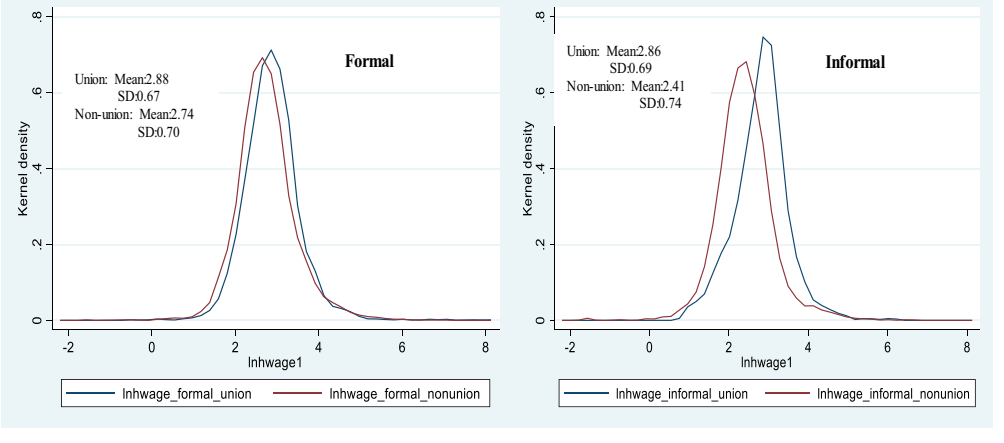

Figure 2. Kernel density of union and nonunion wage distributions among the formal and informal workers.

Figure 2 displays the kernel density of the logarithmic hourly wage distribution by union and nonunion groups among formal and informal workers. The wage gap between union and nonunion workers among informal workers exceeds that among formal workers, suggesting that the union wage premium among informal workers may be larger than that among formal workers.

These descriptive statistical results do not consider factors that may affect wages (e.g. education, occupation, and the industrial sector). The econometric results, controlling for these factors, are discussed in the following sections.

6. Econometric analysis results

6.1. Baseline results of union wage premiums

Table 2 presents the baseline regression results for the wage functions using the FE model. Two estimations are employed. Model 1 estimates for the overall sample using union, employment type (formal worker dummy), demographic factors, work-related factors, region, and survey year FE. Model 2 estimates for the overall sample after adding the interaction term between union and formal workers to Model 1.

First, the results of Model 1 reveal a positive union wage premium (24.6%), significant at the 5% level. The results indicate that the magnitude of China's union wage premium is within the range of some developed countries. For instance, Freeman and Medoff (1984) found a 16% union wage premium in the US, and Farber et al. (2021) reported a range of 10%–30% in developed countries (e.g. the US and UK) from the 1940s to 2010s. This suggests that union functions in China are similar to those in advanced capitalist countries in the recent period. The magnitude of China's union wage premium is broadly consistent with estimates from other developing economies, including South Africa (24% in 2004; Butcher and Rouse 2001), Brazil (15%–28%; Arbache 2009), and India (25% for permanent workers; Bhandari 2008). However, it exceeds the estimates reported for Mexico (3.7%; Rufrancos 2019) and Vietnam (5%–7%; Torm 2014).

Freeman (1980) and Freeman and Medoff (1984) stated that unions are established voluntarily and usually have two faces: a monopoly face of collective bargaining with employers and a voice face of communicating with employees and employers. Both contribute to increasing union members' wage levels. Regarding the union functions in China, the ACFTU is the only legal union organization formed and managed by the government. All unions in the public sector (e.g. SOEs and government organizations)

Table 2. Baseline results of wage functions.

	(1)		(2)	
FE model	Coef.	SE	Coef.	SE
Union	0.246**	0.115	0.607***	0.233
Formal worker	2.037***	0.157	2.125***	0.168
Union × Formal			−0.496**	0.252
Individual fixed effects	Yes		Yes	
Control variable	Yes		Yes	
Year fixed effects	Yes		Yes	
Region fixed effects	Yes		Yes	
Constants	−8.282	7.540	−8.004	7.560
No. of observations	10,414		10,414	
No. of groups	6,745		6,745	
R-squared: Within	0.117		0.118	
Between	0.026		0.031	
Overall	0.036		0.042	
Hausman test: chibar2	191.03		198.43	
Prob > chibar2	0.000		0.000	
F test that all u_i=0	0.000		0.000	

Notes ***$p < 0.01$; **$p < 0.05$; *$p < 0.1$. FE: the fixed-effects model. Control variables include demographic factors (years of schooling, years of work experience and squared term, gender, health status, urban *hukou*, marital status), work-related factors (occupation, workplace ownership type) variables. The estimates for these variables are not presented in the table but can be provided upon request. Standard errors are adjusted using cluster-robust standard error clustering at the individual level.

were managed by the ACFTU during the planned economy period. With market-oriented reforms, labor disputes over wage levels and employment have increased in China since the 1990s. To improve industry relations, the Chinese government has enhanced the functions of the ACFTU to establish unions in the public and private sectors (e.g. POEs and FIEs). Furthermore, since the 2000s, unions have acquired partial rights to engage in collective bargaining with employers, particularly in the private sector (see section 2). The transformation of union functions in China might have increased the influence of unions' collective bargaining agreements on wage setting in China in the 2000s.

Regarding the disparities in the union wage premium across developing countries, unionization has extended into the informal sector in South Africa (Mosoetsa 2001; Goldman 2003) and Latin America (Murillo 2000; Pencavel 1995; Rios-Avila 2017). This expansion might have strengthened the effectiveness of collective bargaining, resulting in an estimated union wage premium of approximately 20%, which is close to the union premium in China. By contrast, the role of unions in Mexico and Vietnam appears to be weaker than that in China, South Africa, and Brazil.

Second, comparing the results of formal and informal workers, the results of Model 2 show that the coefficient of the interaction term between union and formal workers is negative (−0.496) and significant at the 1% level. This suggests that, *ceteris paribus*, the positive effect of unions on the logarithm wages of informal workers is greater than that for formal workers by 49.6% points. Therefore, the union wage premium for informal workers in China is significantly higher than that of formal workers. Union membership among disadvantaged (informal) workers may yield relatively larger wage gains than among advantaged (formal) workers, potentially contributing to narrowing overall wage inequality. These findings are consistent with those of developed countries (D. T. Card 1996; Farber et al. 2021; Freeman 1980; Kerr and Wittenberg 2021).

Third, all results indicate that the coefficients of the formal worker dummy are positive and significant at the 1% level, highlighting a wage gap between formal and informal workers in China.

6.2. Sensitivity checks of union wage premiums

The results of the sensitivity checks are presented in Table 3. All the estimations are performed using the FE model (Eq. [2]). First, we use the yearly wage as the dependent variable (Estimation [a]). Second, as unions mainly protect non-managers, we use samples that exclude managers to rerun the estimations (Estimation [b]). Third, considering the influence of early retirement, we replace the samples aged 16–60 years with those aged 16–50 years (Estimation [c]). Fourth, we include the industry sector and CPC membership variables in Estimation [d].[7] Fifth, we redefine an informal worker as an individual employed in a workplace with fewer than 5 (Estimation [e]) or 10 (Estimation [f]) workers [8]and rerun the models to conduct the sensitivity checks.

The results of Estimations [a]–[f] suggest significant positive union wage premiums ranging from 5.5% to 36.9%. These results are consistent with the findings presented in Table 2.

The interaction terms between the union and formal worker variables are all negative and significant at the 5% or 10% levels in Estimations [a]–[f], suggesting that the union wage premium for informal workers is higher than that for formal workers. This reaffirms the main findings shown in Table 2.

Table 3. Results of sensitivity checks.

	(1)		(2)	
FE model	Coef.	SE	Coef.	SE
[a] Using weekly wage and weekly working hours				
Union	0.055***	0.016	0.109***	0.340
Formal worker	0.067***	0.012	0.074***	0.012
Union × Formal			−0.069**	0.034
Individual fixed effects	Yes		Yes	
Control variables, year and regions fixed effects	Yes		Yes	
No. of observations	19,435		19,435	
No. of groups	11,996		11,996	
[b] Excluding managers				
Union	0.288**	0.118	0.588**	0.248
Formal worker	1.828***	0.165	1.904***	0.178
Union × Formal			−0.417*	0.265
Individual fixed effects	Yes		Yes	
Control variables, year and regions fixed effects	Yes		Yes	
No. of observations	9,353		9,353	
No. of groups	6,202		6,202	
[c] Using samples aged 16–50				
Union	0.369***	0.131	0.864***	0.291
Formal worker	2.233***	0.181	2.328***	0.191
Union × Formal			−0.636**	0.307
Individual fixed effects	Yes		Yes	
Control variables, year and regions fixed effects	Yes		Yes	
No. of observations	8,518		8,518	
No. of groups	5,604		5,604	
[d] Adding industry and party variables				
Union	0.257***	0.139	0.517*	0.307
Formal worker	1.561***	0.227	1.626***	0.247
Union × Formal			−0.330	0.327
Individual fixed effects	Yes		Yes	
Control variables, year and regions fixed effects	Yes		Yes	
No. of observations	5,842		5,842	
No. of groups	4,099		4,099	
[e] Using different definition of informal worker (<5)				
Union	0.231***	0.115	0.677***	0.238
Formal worker	2.130***	0.166	2.243***	0.177
Union × Formal			−0.615**	0.253
Individual fixed effects	Yes		Yes	
Control variables, year and regions fixed effects	Yes		Yes	
No. of observations	10,414		10,414	
No. of groups	6,745		6,745	
[f] Using different definition of informal worker (<10)				
Union	0.231***	0.115	0.643***	0.226
Formal worker	2.005***	0.162	2.116***	0.173
Union × Formal			−0.586**	0.243
Individual fixed effects	Yes		Yes	
Control variables, year and regions fixed effects	Yes		Yes	
No. of observations	10,414		10,414	
No. of groups	6,745		6,745	

Notes ***$p < 0.01$; **$p < 0.05$; *$p < 0.1$. The fixed-effects model is used. The control variables, regional and year fixed effects are the same as those in Table 2. Standard errors are adjusted using cluster-robust standard error clustering at the individual level.

6.3. Robustness checks of union wage premiums

Four methods are employed to conduct the robustness checks for endogeneity issues: Lee's selection-adjusted model (Estimation [a], Eq. [3]), LV method (Estimation [b], Eq. [4]), RE model (Estimation [c], Eq. [2]), and PSM method (Eq. [5]). The results are summarized in Table 4.

Table 4. Results of robustness checks using different models.

	(1)		(2)	
	Coef.	SE	Coef.	SE
[a] Lee selection method				
Union	0.293**	0.092	1.669***	0.228
Formal worker	3.366***	0.090	3.700***	0.100
Union × Formal			−1.890**	0.228
Selection term	3.761***	0.753	4.035***	0.753
Control variables, year and regions fixed effects				
No. of observations	7,109		7,109	
Adjusted R-sq.	0.398		0.403	
[b] LVt-1 model				
Union	0.399***	0.098	0.628***	0.112
Formal worker	3.324***	0.096	3.458***	0.100
Union × Formal			−0.568***	0.097
Control variables, year and regions fixed effects				
No. of observations	6,660		6,660	
Adjusted R-sq.	0.374		0.374	
[c] RE method				
Union	0.313***	0.095	1.290***	0.170
Formal worker	3.035***	0.079	3.243***	0.084
Union × Formal			−1.321***	0.192
Individual fixed effects	Yes		Yes	
Control variable	Yes		Yes	
Year fixed effects	Yes		Yes	
Region fixed effects	Yes		Yes	
No. of observations	10,414		10,414	
No. of groups	6,745		6,745	
R-squared: Within	0.110		0.110	
Between	0.408		0.411	
Overall	0.378		0.382	
BP test: chibar2(01)	677.410		648.620	
Prob > chibar2	0.000		0.000	

Notes ***$p < 0.01$; **$p < 0.05$; *$p < 0.1$. LV *t*-1: the model with lagged term (prior survey waves) of union dummy variable is used; RE: random effects model; BP test: the Breusch-Pagan Lagrange multiplier (LM) test for the random effects model. The control variables, regional and year fixed effects are the same as those in Table 2. Standard errors are adjusted using cluster-robust standard error clustering at the individual level.

First, the results of Estimations [a]–[c] suggest significantly positive union wage premiums ranging from 29.3% to 39.9%. These results are consistent with the findings presented in Table 2. The interaction term between union and formal workers are negative and significant at the 1% level in Estimations [a]–[c]. This suggests that informal workers' union wage premium is higher than that of formal workers, which reaffirms the main findings shown in Table 2.

Second, the coefficients of the selection-adjusted terms in Estimation [a] are positive and significant at the 1% level, suggesting that the results may be underestimated if the sample selection bias issue is not addressed.

Third, Table 5 summarizes the results obtained using several PSM methods, including the one-to-one nearest neighbor (Model [a]), k-nearest neighbor (1:3 matching in this study; Model [b]), kernel (Model [c]), radius (Model [d]), and local linear regression (Model [e]) matching methods. Figure 3 shows the standard bias across the observable variables before and after matching the total sample in a balance check. The differences in the standard bias of covariates between union members and non-members became smaller (approaching a value near '0') after matching. Figure 4 shows the results of the overlap check. These results indicate that the matching procedures are appropriate.

Table 5. Results of wage union premium among overall, formal and informal workers based on the psm method.

	Sample	Treated	Controls	Difference	SE	t-value
[a] One to one Nearest-neighbor e matching (1:1)						
Total	Unmatched	2.841	2.573	0.267	0.021	12.57
	ATT	2.833	2.775	0.058	0.037	1.55
Formal	Unmatched	2.855	2.716	0.139	0.024	5.82
	ATT	2.847	2.832	0.016	0.038	0.41
Informal	Unmatched	2.774	2.382	0.391	0.048	8.24
	ATT	2.768	2.660	0.108	0.092	1.17
[b] k-Nearest-neighbor matching (1:3)						
Total	Unmatched	2.841	2.573	0.267	0.021	12.57
	ATT	2.833	2.793	0.040	0.031	1.30
Formal	Unmatched	2.855	2.716	0.139	0.024	5.82
	ATT	2.847	2.829	0.018	0.033	0.56
Informal	Unmatched	2.774	2.382	0.391	0.048	8.24
	ATT	2.768	2.644	0.124	0.074	1.69
[c] Kernel matching						
Total	Unmatched	2.841	2.573	0.267	0.021	12.57
	ATT	2.841	2.785	0.056	0.028	1.98
Formal	Unmatched	2.855	2.716	0.139	0.024	5.82
	ATT	2.854	2.811	0.044	0.031	1.41
Informal	Unmatched	2.774	2.382	0.391	0.048	8.24
	ATT	2.775	2.681	0.094	0.067	1.41
[d] Radius matching						
Total	Unmatched	2.841	2.573	0.267	0.021	12.57
	ATT	2.835	2.786	0.049	0.028	1.74
Primary	Unmatched	2.855	2.716	0.139	0.024	5.82
	ATT	2.848	2.813	0.035	0.031	1.13
Secondary	Unmatched	2.774	2.382	0.391	0.048	8.24
	ATT	2.750	2.640	0.109	0.068	1.60
[e] Local linear regression matching						
Total	Unmatched	2.841	2.573	0.267	0.021	12.57
	ATT	2.833	2.789	0.044	0.037	1.18
Formal	Unmatched	2.855	2.716	0.139	0.024	5.82
	ATT	2.847	2.819	0.028	0.038	0.74
Informal	Unmatched	2.774	2.382	0.391	0.048	8.24
	ATT	2.768	2.676	0.092	0.092	0.99

Notes: The logit regression model is used to estimate the probability score of becoming a formal worker, The covariate variables include demographic factors (years of schooling, years of work experience and squared term, gender, health status, urban *hukou*, CPC membership), work-related factors (occupation, industry sector, workplace ownership type), family background factors (married, father's education, mother's education, father with CPC membership, mother with CPC membership), region, and year dummies. The variables in wage functions are the same as those in Table 2.

The results of the five matching methods are consistent. For example, based on the results of the one-to-one nearest neighbor matching (Model [a]), the differences in wage levels between the treatment (union member) and control (nonunion member) groups for the matched and unmatched groups exhibit positive values. However, magnitude of the differences in the matched group (0.058) is smaller than that in the unmatched group (0.267). This finding suggests that individual heterogeneity may affect the probability of becoming a union member, which significantly affects the union wage premium. Furthermore, the magnitude of the ATT values for informal workers (0.108) is greater than that for formal workers (0.016). These results confirm the findings provided in Table 2.

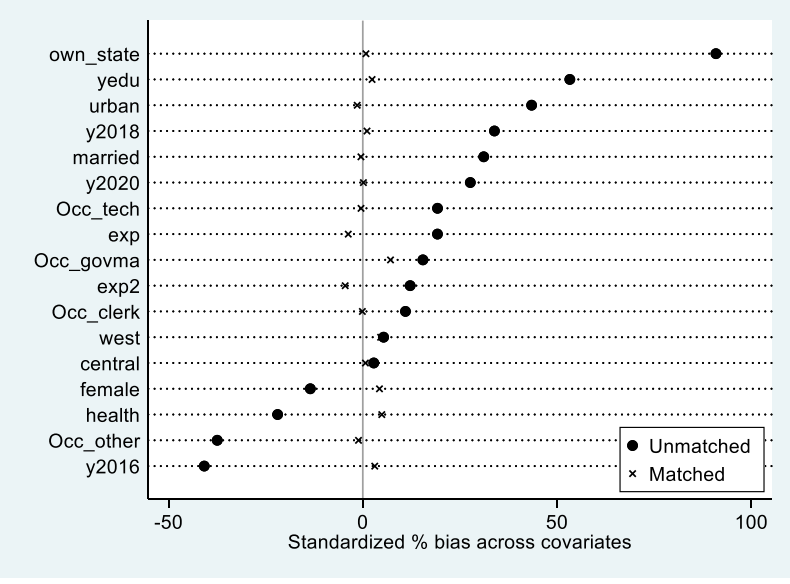

Figure 3. Standardized bias across variables among unmatched and matched group (overall samples).

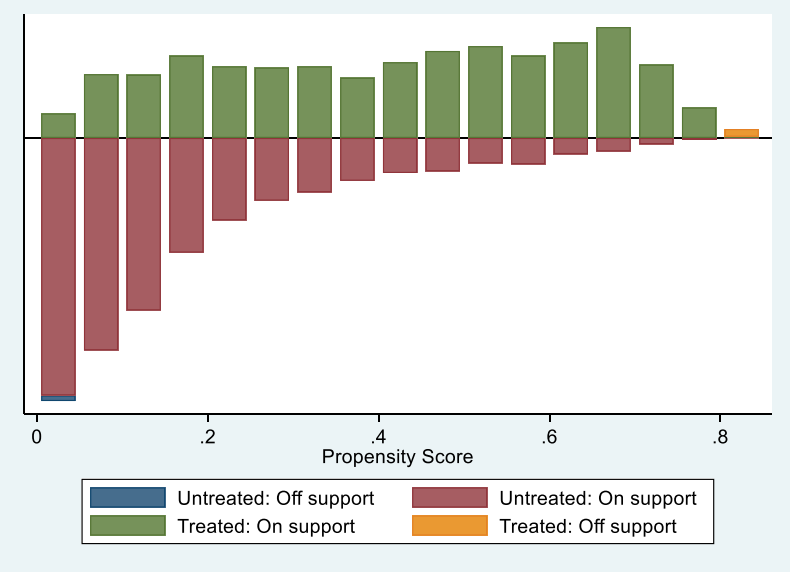

Figure 4. Results of overlap checks (overall samples).

6.4. Baseline decomposition results

Table 6 presents the decomposition results for the wage gap between formal and informal workers based on the basic B-O decomposition method. Models 1 and 2 employ Eqs. (6) and (7), respectively.

Table 6. Decomposition results based on the basic B-O decomposition method.

	(1)		(2)	
B-O Method and RE	Endowment	Price	Endowment	Price
Value: Total wage gap=0.339	0.192	0.147	0.182	0.157
Contribution rate (%)				
Total	56.6%	43.4%	53.7%	46.3%
Each factor				
Union	**2.3%**	**-2.8%**	**6.8%**	**-7.2%**
Education	40.9%	104.7%	22.9%	122.7%
Experience	−5.5%	23.0%	−1.8%	19.3%
Gender	2.5%	14.2%	3.9%	12.8%
Urban *hukou*	−0.2%	−9.3%	0.6%	−10.1%
Married	0.1%	−15.5%	−0.7%	−14.7%
Health	0.1%	−10.6%	−0.5%	−10.0%
Occupation	12.8%	8.7%	4.6%	16.8%
Public sector	−5.4%	−17.8%	8.7%	−32.0%
Year	−0.1%	−10.9%	2.2%	−13.2%
Region	9.2%	−10.3%	7.0%	−8.1%
Constant	0.0%	−30.1%	0.0%	−30.0%

Notes: The results obtained from the basic B-O decomposition method based on the wage function using the random-effects method. Model 1 is the decomposition based on Eq. (6), Model 2 is the decomposition based on Eq. (7) described in the text. The results of wage functions are presented in Appendix Table A2.

First, both the endowment and price effect components contribute to the formation of the wage gap. In Models 1 and 2, the endowment effect is larger (56.6% and 53.7%, respectively) than the price effect (43.4% and 46.3%, respectively). Thus, wage gaps are generated by disparities in the endowment and price effects between formal and informal workers. Notably, the endowment effect exceeds the price effect.

Second, regarding the effects of the union on the wage gap, the results indicate that the disparity in union density (2.3% and 6.8%) widens the wage gap, whereas the difference in union wage premiums (−2.8% and −7.2%) narrows the wage gap.

Third, other factors influence the formation of the wage gap. The contributions of the years of schooling (40.9% and 22.9%) and returns on education (104.7% and 122.7%) to the wage gap are the greatest. Thus, education is the primary factor contributing to the formation of the wage gap. The second most influential factors for the endowment effect are occupation (Model 1) and ownership sector (Model 2), and the second most influential factor for the price effect is years of work experience (Models 1 and 2). This suggests that occupational segregation, labor market segmentation into the public and private sectors, and differences in returns on human capital significantly influence the wage gap. These results are consistent with the human capital theory and labor market segmentation hypothesis.

Additionally, based on the informal worker wage function in the decomposition method (Column (2) in Appendix Table A2), when an informal worker changes from a nonunion member to a union member, the logarithm wage increases by 11.8% points. Furthermore, when an informal worker's years of schooling increase by one, the wage rises by 3.8% points. These findings suggest that, besides the human capital factors, union membership plays an important role in raising informal workers' wage levels, potentially narrowing overall wage inequality.

Table 7. Robustness checks of decompositions.

	(1) Q-R decomposition and RE				(2) Q-R decomposition and RE			
	Endowment		Price		Endowment	Price	Endowment	Price
	(a)	(b) loss of non-union	(c) gain of union	(d) total (b+c)	(e)	(f)	(g)	(h)
Value	0.214	0.085	0.040	0.125	0.191	0.141	0.189	0.143
Contribution rate (%)								
Total	**63.2%**	**25.0%**	**11.8%**	**36.8%**	**59.6%**	**40.4%**	**59.0%**	**41.0%**
Each factor								
Union	**3.9%**	**-1.8%**	**-2.6%**	**-4.4%**	**1.8%**	**-3.7%**	**8.2%**	**-10.1%**
Education	35.6%	73.7%	36.3%	110.0%	38.0%	82.5%	23.0%	97.5%
Experience	−3.0%	5.0%	15.5%	20.5%	−5.2%	36.2%	−0.2%	31.4%
Health	3.1%	8.1%	5.5%	13.6%	2.8%	13.6%	4.4%	12.1%
Gender	0.1%	−5.0%	−4.7%	−9.7%	0.0%	3.7%	−0.5%	4.2%
Ethnicity	−0.3%	−7.7%	−7.4%	−15.1%	0.4%	−35.9%	−1.1%	−34.4%
Married	−0.1%	−7.1%	−3.3%	−10.4%	1.1%	−3.1%	2.1%	−4.1%
Occupation	13.7%	6.0%	1.8%	7.6%	12.5%	3.3%	1.6%	14.3%
Public sector	−0.4%	−11.5%	−11.3%	−22.8%	−5.9%	−14.5%	7.6%	−27.9%
Region	1.4%	−7.6%	−4.9%	−12.5%	0.5%	−4.2%	2.1%	−5.9%
Year	9.2%	−10.3%	−0.1%	−10.4%	10.9%	−7.7%	9.0%	−6.0%
Party					0.1%	−12.0%	−0.8%	−11.1%
Industry					2.4%	−1.3%	3.5%	−2.5%
Constant	0.0%	−16.8%	−13.1%	−29.6%	0.0%	−16.3%	0.0%	−16.3%

Notes: Model1: the Oaxaca-Ransom decomposition method based on the wage functions using the RE model; Model2: adding industrial sector and party membership factors to Model1.

6.5. Robustness checks for the decomposition results

We use two methods to conduct robustness checks for the decomposition analyses (Table 7). First, to address the index number problem in the basic B-O decomposition method (Eqs. [6] and [7]), we employ the O-R decomposition method based on wage functions using the RE model (Eq. [8]). The endowment and price effects of the union are 3.9% (Column [a]) and −4.4% (Column [d]), respectively. These findings are consistent with those presented in Table 6.

Second, we extend the factors by adding the industrial sector and party membership variables. The O-R decomposition method based on wage functions using the RE model is used. The endowment and price effects of the union are 1.8% (Column [e]) and −10.1% (Column [h]), respectively, reaffirming the main findings in Table 6.

7. Conclusion

This study explores the impact of unions on the wage gap between formal and informal workers in China based on the national longitudinal survey data from the CFPS from 2014 to 2020. The results demonstrate that while a positive union wage premium exists for both formal and informal workers, the premium for informal workers is larger than that for formal workers. We found that, in China, the disparity in union membership density contributes to widening the wage gap, whereas the difference in union wage premiums contributes to narrowing the wage gap.

The findings of this study have several practical implications. First, the union wage premium in China is within the range of the premiums in developed countries (Rosenfeld and Denice 2019; Choi and Ramos 2021; Farber et al. 2021; Kerr and Wittenberg 2021). The results suggest that, with the economic system transitioning from a planned economy to a market-oriented economy, the influence of unions on wage settings has increased significantly. The collective bargaining agreement has recently led to positive union premiums in China. This may be because of the development of the private sector in China. Although Chinese unions in the public sector (e.g. SOEs) are still managed by the government (the ACFTU), unions in the private sector (e.g. POEs and FIEs) may influence wage setting through collective bargaining agreements, such as unions in advanced capitalist countries.

Second, the decomposition results indicate that disparities in union membership density widens the wage gap. Based on the CFPS, the proportion of union members among formal and informal workers from 2014 to 2020 was 18.8%–40.4% and 8.7%–24.8%, respectively. Most informal workers are in the private sector, where the union coverage rate is low. Following the *Trade Union Law of the People's Republic of China*, employers in the private sector can choose to establish unions in their workplaces. However, most private sector employers do not permit the establishment of unions to avoid government interference. Consequently, the union coverage rate in the private sector is lower than in the public sector. The expansion of union coverage in the private sector, particularly by incorporating informal workers, potentially narrows the wage gap between formal and informal workers. As the share of informal employment is substantially larger in developing economies, including China, extending unionization to the informal sector is expected to enhance working conditions, including wage growth and employment security. Such improvements are likely to reduce income inequality and foster sustainable development in developing countries.

However, the expansion of union coverage does not necessarily translate into a uniform reduction in wage inequality. Workers may choose to enter firms with already established unions. Beauregard et al. (2025) found that high-productivity firms are more likely to be unionized than those with lower productivity. In the Chinese context, while the ACFTU has created union branches in the public sector that cover virtually all formal and most informal workers, union coverage in the private sector remains comparatively limited. Broadening union coverage (or increasing the union density) raises wages primarily for workers in more productive firms and may widen the wage gap between high- and low-productivity firms in the private sector.

Finally, this study has several limitations. Despite employing the FE, LV, RE, and PSM models to mitigate endogeneity concerns, establishing a causal link between unionization and the wage gap requires further investigation. The household survey data used in this study provide limited information on workplace characteristics, preventing an analysis of firm-level factors such as productivity. Future research should integrate employer – employee matched data to examine these dimensions. Distinguishing the effects of unions on wage differentials within and between establishments can enhance our understanding of the relationship between unionization and wage inequality, particularly in the context of self-selection. Furthermore, examining variations in union effects across regions or industry sectors could offer important insights into the nexus between economic development, unionization, and income inequality.

Nonetheless, this study employs recent nationally representative longitudinal survey data to provide novel evidence on the impact of unions on the wage gap between formal and informal workers in China. Given the increasing incidence of informal employment and the extension of unionization to informal workers in other developing economies, such as South Africa (Goldin 2003), Latin America (Rios-Avila 2017), and India (Bhandari 2008), our empirical findings offer broader insights into the role of unions in the labor markets of developing economies.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The work was supported by the Japan Society for the Promotion of Science [120257301].

ORCID

Xinxin Ma http://orcid.org/0000-0001-7616-1941

References

Abraham, F., J. Konings, and S. Vanormelingen. 2009. "The Effect of Globalization on Union Bargaining and Price-Cost Margins of Firms." *Review of World Economics* 145 (1): 13–36. https://doi.org/10.1007/s10290-009-0003-8.

Arias, J., E. Artuc, D. Lederman, and D. Rojas. 2018. "Trade, Informal Employment and Labor Adjustment Costs." *Journal of Development Economics* 133:396–414. https://doi.org/10.1016/j.jdeveco.2018.03.006.

Azam, J., and S. Rospabe. 2007. "Trade Unions vs Statistical Discrimination: Theory and Application to Post-Apartheid South Africa." *Journal of Development Economics* 84 (1): 417–444. https://doi.org/10.1016/j.jdeveco.2005.12.005.

Banerjee, A., S. Galiani, J. Levinsohn, Z. Mclaren, and I. Woolard. 2008. "Why Has Unemployment Risen in the New South Africa." *The Economics of Transition* 16 (4): 715–740. https://doi.org/10.1111/j.1468-0351.2008.00340.x.

Beauregard, P. L., B. . T. Lemieux, D. Messacar, and R. Saggio. 2025. "Why Do Union Jobs Pay More? New Evidence From Matched Employer-Employee Data." Working Paper No.33740, National Bureau of Economic Research.

Becker, G. S. 1957. *The Economics of Discrimination*. Chicago, Illinois: The University of Chicago Press.

Becker, G. S. 1964. *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. New York: Columbia University Press.

Bhandari, A. K. 2008. "Union Membership Effect on Wage Premiums: Evidence from Organized Manufacturing Industries in India." IZA discussion paper no. 3747.

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

Butcher, K. F., and C. E. Rouse. 2001. "Wage Effects of Unions and Industrial Council in South Africa." *Industrial & Labor Relations Review* 54 (2): 349–374.

Card, D. 1996. "The Effect of Unions on the Structure of Wages: A Longitudinal Analysis." *Econometrica: Journal of the Econometric Society* 64 (4): 957–979. https://doi.org/10.2307/2171852.

Card, D., T. Lemieux, and W. C. Riddell. 2020. "Unions and Wage Inequality: The Roles of Gender, Skill and Public Sector Employment." *Canadian Journal of Economics/Revue Canadienne d'économique* 53 (1): 140–173. https://doi.org/10.1111/caje.12432.

Chang, J., and H. Huang. 2016. "Trade Unions, Unemployment, Economic Growth, and Income Inequality." *Macroeconomic Dynamics* 20 (1): 404–428. https://doi.org/10.1017/S1365100514000443.

Chattaraj, S. 2016. "Organising the Norganised Union Membership and Earning in India's Informal Economy." BSG Working Paper Series No.015.

Choi, H., and R. Ramos. 2021. "The Union Wage Premium in a Segmented Labor Market: New Evidence from Korea." *Journal of the Asia Pacific Economy* 28 (1): 242–260. https://doi.org/10.1080/13547860.2021.1879352.

Clarke, S., C. H. Lee, and D. Q. Chi. 2007. "From Rights to Interests: The Challenge of Industrial Relations in Vietnam." *Journal of Industrial Relations* 49 (4): 545–568. https://doi.org/10.1177/0022185607080321.

Dickens, W. T., and K. Lang. 1985. "A Test of Dual Labor Market Theory." *American Economic Reviews* 75 (4): 792–805.

DiNardo, J., N. M. Fortin, and T. Lemieux. 1996. "Labor Markets Institutions and the Distribution of Wages, 1973–1992: A Semiparametric Approach." *Econometrica* 64 (5): 1001–1044. https://doi.org/10.2307/2171954.

Farber, H. S., D. Herbst, I. Kuziemko, and S. Naidu. 2021. "Unions and Inequality over the Twentieth Century: New Evidence from Survey Data." *Quarterly Journal of Economics* 136 (3): 1325–1385. https://doi.org/10.1093/qje/qjab012.

Fichtenbaum, R. 2006. "Labour Market Segmentation and Union Wage Gaps." *Review of Social Economy* 64 (3): 387–420. https://doi.org/10.1080/00346760600892808.

Freeman, R. B. 1980. "The Exit-Voice Tradeoff in the Labor Market: Unionism, Job Tenure, Quits and Separations." *Quarterly Journal of Economics* 94 (4): 643–673.

Freeman, R. B. 2009. "Labor Regulations, Unions, and Social Protection in Developing Countries: Market Distortions or Efficient Institutions?" NBER Working Paper Series No. 14789.

Freeman, R. B., and J. L. Medoff. 1984. *What Do Unions Do?* New York: Basic Books.

Ge, Y. Y., C. P. Li, and Y. Ge. 2022. "The Effect of Informal Employment and Educational Mismatch on Wages of Industrial Workers." *Finance & Economics* 3:123–134 (In Chinese).

Goldman, T. 2003. "Organising in South Africa's Informal Economy: An Overview of Four Sectoral Case Studies." *In Focus Programme on Boosting Employment Through Small Enterprise Development*. Geneva: International Labor Organization.

Gunderson, M., B. Y. Lee, and H. Wang. 2016. "Union Pay Premium in China: An Individual-Level Analysis." *International Journal of Manpower* 37 (4): 606–627. https://doi.org/10.1108/IJM-08-2014-0174.

Guo, H., and W. Dai. 2022. "Original Aspiration and Mission: Historical Review and Experience of Chinese Trade Unions Participating in National Governance." *Journal of Beijing Federation of Trade Unions Cadre College* 37 (4): 53–62 (In Chinese).

Herrera-Idárraga, P., E. López-Bazo, and E. Motellón. 2015. "Double Penalty in Returns to Education: Informality and Educational Mismatch in the Colombian Labour Market." *The Journal of Development Studies* 51 (12): 1683–1701. https://doi.org/10.1080/00220388.2015.1041516.

International Labor Organization. 2015. *R204–Transition from the Informal to the Formal Economy Recommendation: Recommendation Concerning the Transition from the Informal to the Formal Economy*. Geneva: International Labor Organization.

International Labor Organization. 2018. *Women and Men in the Informal Economy: A Statistical Picture* (3rd ed.). Geneva: International Labor Organization.

Jolevski, F., G. Nayyar, R. Pleninger, and S. Yu. 2025. "Spillovers in ICT Adoption from Formal to Informal Firms: Evidence from Zambia." *Journal of Development Economics* 177:103549. https://doi.org/10.1016/j.jdeveco.2025.103549.

Kerr, A., and M. Wittenberg. 2021. "Union Wage Premia and Wage Inequality in South Africa." *Economic Modelling* 97:255–271. https://doi.org/10.1016/j.econmod.2020.12.005.

Kim, M., and G. Kim. 2021. "Effects of Korean Trade Union on Wage Discrimination Between Regular and Irregular Workers." *Asian Economic Journal* 37 (1): 3–14.
Lee, L. F. 1983. "Generalized Econometric Models with Selectivity." *Econometrica* 51 (2): 507–512. https://doi.org/10.2307/1912003.
Lewis, H. G. 1963. *Unionism and Relative Wages in the United States*. The University of Chicago Press.
Lewis, W. A. 1954. "Economic Development with Unlimited Supplies of Labor." *The Manchester School of Economic and Social Studies* 28 (2): 139–151.
Li, M. S., and X. Gao. 2023. "Difference in Returns to Schooling Between Formal and Informal Employment in China." *Applied Economics* 55 (18): 2098–2111. https://doi.org/10.1080/00036846.2022.2101610.
Li, M., and J. Xu. 2014. "Who Benefits from China's Trade Union." *Economic Research Journal* 5:49–61. (In Chinese).
Lin, K. J., X. Lu, J. Zhang, and Y. Zheng. 2020. "State-Owned Enterprises in China: A Review of 40 Years of Research and Practice." *China Journal of Accounting Research* 13 (1): 31–55. https://doi.org/10.1016/j.cjar.2019.12.001.
Mosoetsa, S. 2001. "The Manchester Road: Women and the Informalisation of Work in South Africa's Footwear Industry." *Labour, Capital & Society* 34 (2): 184–207.
Murillo, M. V. 2000. "From Populism to Neoliberalism: Labor Unions and Market Reforms in Latin America." *World Politics* 52 (2): 135–168. https://doi.org/10.1017/S0043887100002586.
Ntlhola, M. A., P. Kwenda, and M. Ntuli. 2019. "A Distributional Analysis of Union-Wage Effects in South Africa: Evidence from Panel Data." *Development Southern Africa* 36 (3): 309–328. https://doi.org/10.1080/0376835X.2018.1484700.
Oaxaca, R. L. 1973. "Male-Female Wage Differentials in Urban Labor Markets." *International Economic Review* 14 (3): 693–709. https://doi.org/10.2307/2525981.
Oaxaca, R. L., and M. R. Ransom. 1994. "On Discrimination and the Decomposition of Wage Differentials." *Journal of Econometrics* 61 (1): 5–21. https://doi.org/10.1016/0304-4076(94)90074-4.
Oberfichtner, M., C. Schnabel, and M. Töpfer. 2020. "Do Unions and Works Councils Really Dampen the Gender Pay Gap? Discordant Evidence from Germany." *Economics Letters*196: article 109509. *Economics Letters* 196:109509. https://doi.org/10.1016/j.econlet.2020.109509.
Pencavel, J. 1995. "The Role of Labor Unions in Fostering Economic Development." Policy, Research Working Paper No. 1469, Washington, D.C: World Bank Group.
Ramoni-Perazzi, J., and G. Orlandoni-Merli. 2021. "Analysis of the Formal/Informal Wage Inequalities in Colombia: A Semiparametric Approach." *Journal of Applied Social Science* 15 (1): 107–131. https://doi.org/10.1177/1936724420975343.
Rios-Avila, F. 2017. "Unions and Economic Performance in Developing Countries: Case Studies from Latin America." *Ecos de Economía* 21 (44): 4–36. https://doi.org/10.17230/ecos.2017.44.1.
Rosenbaum, P. R., and D. B. Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 70 (1): 41–55. https://doi.org/10.1093/biomet/70.1.41.
Rosenfeld, J., and P. Denice. 2019. "What Do Government Unions Do? Public Sector Unions and Nonunion Wages, 1977–2015." *Social Science Research* 78:41–56. https://doi.org/10.1016/j.ssresearch.2018.10.011.
Rufrancos, H. G. 2019. "Are There Gains to Joining a Union? Evidence from Mexico." *British Journal of Industrial Relations* 57 (3): 676–712. https://doi.org/10.1111/bjir.12463.
Shahen, M. E., K. Kotani, M. Kakinaka, and S. Managi. 2020. "Wage and Labor Mobility Between Public, Formal Private and Informal Private Sectors in a Developing Country." *Economic Analysis and Policy* 68:101–113. https://doi.org/10.1016/j.eap.2020.09.006.
Sun, Z., and Y. Liu. 2019. "How Does Trade Union Influence Workers' Wage in China: An Analysis Based on CGSS 2008-2015 Mixed Cross-Sectional Data." *Labor Economic Studies* 7 (4): 121–144 (In Chinese).
Todaro, M. P. 1969. "A Model of Migration and Urban Unemployment in Less Developed Countries." *The American Economic Review* 59 (1): 138–148.
Torm, N. 2014. "The Role of Trade Unions in Vietnam: A Case Study of Small and Medium Enterprises." *Journal of International Development* 26 (2): 207–221. https://doi.org/10.1002/jid.2881.

Yao, Y., and N. Zhong. 2013. "Unions and Workers' Welfare in Chinese Firms." *Journal of Labor Economics* 31 (3): 633–667. https://doi.org/10.1086/669819.

You, Z. 2017. "On Some Quantitative Studies into Labor Union in China." *Journal of Southeast University (Philosophy and Social Science)* 19 (5): 60–66 (In Chinese).

Zhang, K. S., C. H. Liu, and S. L. Ding. 2018. "Research on the Wage Differences Between Formal Employment and Informal Employment." *Chinese Journal of Population Science* 1:83–94 (In Chinese).

Zhang, Y. J., and B. Qin. 2015. "Research on the Wage Gap Between Urban Informal Labor and Formal Labor." *Population Journal* 4:92–103 (In Chinese).

Appendix

Table A1. Definition and descriptive statistics of variables (overall samples).

	Definitions	Mean	SD
Lnwage	logarithm of the hourly wage calculated based on total wage and work hours	2.646	0.720
Union	1=trade union member, 0=non-member	0.199	0.399
Formal employee	1=formal employee, 0= informal employee	0.503	0.500
Education	Years of schooling	12.571	3.997
Experience	Years of experience=age-6-years of schooling	20.751	12.832
Female	1=women, 0=men	0.456	0.498
Urban *hukou*	1=owning urban *hukou*, 0=otherwise	0.881	0.324
Married	1=having a spouse, 0=otherwise	0.771	0.420
Health	1=health status is very good or good, 0=otherwise	0.326	0.469
Occupation			
Manager	1=manager or technician, 0=otherwise	0.102	0.303
Technician	1=technician, 0=otherwise	0.189	0.391
Clerk	1=clerk, 0=otherwise	0.143	0.350
Operator	1=operator, 0=otherwise	0.231	0.421
Other occupation	1=other occupation, 0=otherwise	0.336	0.472
Public sector	1=public sector, 0=otherwise	0.334	0.472
Region			
East	1=eastern reagion, 1=otherwise	0.492	0.500
Central	1=central region, 2=otherwise	0.331	0.471
West	1=western region, 3=otherwise	0.177	0.382
Year			
y2014	1=2014 survey year, 0=otherwise	0.316	0.465
y2016	1=2016 survey year, 0=otherwise	0.196	0.397
y2018	1=2018 survey year, 0=otherwise	0.352	0.478
y2020	1=2020 survey year, 1=otherwise	0.137	0.343
Party	1= a member of Communist Party of China, 0=non-CPC member	0.164	0.370
Industry			
Manufacturing and construction	1=working in the manufactural or construction industry sector, 0=otherwise	0.302	0.459
Traffic and communication	1=working in the traffic or communication industry sector, 0=otherwise	0.058	0.234
Retail and catering	1=working in the traffic or sale industry sector, 0=otherwise	0.197	0.397
Service	1=working in the traffic or service industry sector, 0=otherwise	0.292	0.455
Other industry	1=other indusrty, 0=otherwise	0.152	0.359

Notes: The samples were limited to those aged 16-59.

Table A2. Results of wage functions by formal and informal workers.

	(1) Formal		(2) Informal	
	Coef.	SE	Coef.	SE
Union	0.040 **	0.021	0.118 ***	0.041
Education	0.068 ***	0.005	0.038 ***	0.006
Exp.	0.013 ***	0.003	0.008 *	0.005
Exp._sq.	0.000 **	0.000	0.000 ***	0.000
Female	-0.189 ***	0.021	-0.288 ***	0.030
Urban *hukou*	-0.011	0.048	0.026	0.050
Married	-0.006	0.027	0.062	0.039
Health	-0.012	0.019	0.094 ***	0.027
Occupation (Ref. Operator)				
Manager	0.214 ***	0.035	-0.091	0.104
Technician	0.211 ***	0.030	0.173 ***	0.049
Clerk	0.002	0.030	-0.057	0.046
Other occupation	-0.002	0.028	-0.023	0.035
Public sector	-0.081 ***	0.021	0.130 ***	0.036
Region (Ref. East)				
Central	-0.254 ***	0.023	-0.193 ***	0.031
West	-0.240 ***	0.028	-0.177 ***	0.040
Year (Ref. 2014)				
y2016	0.121 ***	0.035	0.140 ***	0.040
y2018	0.084 ***	0.020	0.162 ***	0.029
y2020	0.142 ***	0.025	0.211 ***	0.038
Constants	1.783 ***	0.083	1.894 ***	0.096
No. of observations	5242		5172	
No. of groups	3,495		3,860	
R-squared: Within	0.083		0.081	
Between	0.208		0.159	
Overall	0.201		0.170	
BP test: chibar2	319.69		97.41	
Prob > chibar2	0.000		0.000	

Notes: *** $p<0.01$; ** $p<0.05$; * $p<0.1$. The random effects model was used; BP test: the Breusch-Pagan Lagrange multiplier (LM) test for the random effects model. Standard errors are adjusted using cluster-robust standard error clustering at the individual level.