Ma *J Labour Market Res* (2024) 58:3
https://doi.org/10.1186/s12651-024-00361-2

Journal for Labour Market Research

ORIGINAL ARTICLE

Open Access

Union membership and the wage gap between the public and private sectors: evidence from China

Xinxin Ma[1*]

Abstract

As trade unions are active in corporations worldwide, their effects on the labor market have attracted global attention. However, there is scarce empirical evidence regarding how trade unions' effects on the wage gap differ between the public and private sectors, especially in the Chinese context. Using national longitudinal survey data from the China Family Panel Studies for the years 2010, 2012, 2014, 2016, and 2018 and the Blinder-Oaxaca decomposition method, this study estimates unions' effect on the wage gap between the public and private sectors in China. The results from the Ordinary Least Squares (OLS) method indicate the existence of a significant positive union wage premium nationwide. Additionally, the premium in the public sector is greater than that in the private sector. However, this effect becomes insignificant after accounting for individual heterogeneity using the fixed effects model. The decomposition results based on the OLS method indicate that the union coverage difference (the endowment effect) widens the wage gap between the public and private sectors; conversely, the union wage premium difference (the price effect) narrows the wage gap. These results indicate that a policy expanding union coverage in the private sector may effectively narrow the wage gap between both sectors.

Keywords Wage gap, Public and private sector, Union coverage, Union wage premium, China

JEL Classification J51, J53, J31, P21, P31, P41

1 Introduction

As trade unions (hereinafter, unions) are active in corporations worldwide, their effects on the labor market have attracted global attention in developed (Lewis 1963; Freeman 1980; Parsley 1980; Hirsch and John 1986; Card 1996; Chang and Huang 2016; Farber et al. 2021; Kulkarni and Hirsch 2021; Masso et al. 2022), developing (Casale and Posel 2010; Gunderson 2016; Kerr and Wittenberg 2021), and emerging market economies (Yuan 2015; Magda et al. 2016; Li and Song 2017; Yang et al. 2018; Sun and Liu 2019). A union can protect its members through collective bargaining and by obtaining rent from a corporation. Consequently, this would lead to an increase in the union wage premium (e.g., Lewis 1963; Freeman 1980). Higher wages for union members can lead to a wage gap or income inequality (Card 1996). Numerous studies have examined the "union wage premium" in developed countries (Lewis 1963; Freeman 1980; Card 1996; Blanchflower and Bryson 2010; Rosenfeld and Denice 2019; Choi and Ramos 2021; Kerr and Wittenberg 2021) and developing or emerging market economies, including China (Yao and Zhong 2013; Li and Xu 2014; Yuan 2015; Gunderson et al. 2016; Magda et al. 2016; Li and Song 2017; Yang et al. 2018; Sun and Liu 2019). However, there is scarce empirical evidence regarding how unions' effects on the wage gap differ between the public and private sectors (Blanchflower and Bryson 2010; Rosenfeld and Denice 2019). This study is the first to

*Correspondence:
Xinxin Ma
xxma@hosei.ac.jp
[1] Faculty of Economics, Hosei University, 4342 Machita-Shi Aiharamachi, Tokyo 194-0298, Japan

investigate union membership's effects on the wage gap between the public and private sectors in China, the largest country undergoing an economic system transition.

The wage gap between the public and private sectors is a hot topic in both developed and emerging market economies (Cardozo and Cunha 2019; Gindling et al. 2020; Maria and Christofides 2020; Sławinska 2020; Ma and Li 2022). Numerous studies have argued that because the public sector [e.g., state-owned enterprises (SOEs)] is usually supported by the government, it can easily become a monopoly sector that obtains a higher monopoly or political rent (Shleifer and Vishny 1994, 1997; Lin et al. 2020; Iwasaki et al. 2022; Jin et al. 2022). Consequently, it can set a higher wage level than that in the private sector, which leads to a wage gap between both sectors.

Many countries have aimed to narrow the wage gap between the public and private sectors, yet this issue persists worldwide (Cardozo and Cunha 2019; Gindling et al. 2020; Maria and Christofides 2020; Sławinska 2020; Ma and Li 2022). Using wage decomposition methods, Chen et al. (2005), Zhang and Xue (2008), Ye et al. (2011), Demurger et al. (2012), Zhang (2012), Ma (2018), and Ma and Li (2022) found that differences in workers' endowment (e.g., education) and wage-setting systems (e.g., return to education or work experience) generate a wage gap between the public and private sectors. However, prior studies have not investigated unions' effects on that gap.

Blanchflower and Bryson (2010) and Rosenfeld and Denice (2019) reported that both union membership density, collective agreement coverage and union wage premiums differences exist between the public and private sectors in the US and the UK. However, they have not investigated how the two components—(i) the difference in union coverage (the endowment effect) and (ii) the difference in union wage premium (the price effect)—affect the wage gap simultaneously. Further, policy implications differ between the two components: when the difference in union coverage is the main component, expanding union coverage in the private sector is expected to narrow the wage gap; when the difference in the union wage premium is the main component, the policy of improving the collective bargaining function in the private sector may contribute to narrowing the wage gap. Therefore, both components' magnitudes must be examined simultaneously and compared. Thus, a concurrent investigation of each component's magnitude has both academic and practical significance. The purpose of this study is to provide some empirical evidence to clarify both components, which could offer valuable insights for policymaking.

I chose China as the research setting for two reasons. First, the proportion of workers in the public sector is inconsequential in developed countries. In contrast, China institutionally transitioned from a planned economy to a market-oriented economy in 1978. Furthermore, as its government implemented a gradualist reform strategy aiming to maintain the existence of large-size SOEs (Lin et al. 1994, 2020; Iwasaki et al. 2022; Jin et al. 2022), China has the greatest proportion of workers in the public sector worldwide. Hence, a sufficiently large sample is available to fully examine this issue. Second, because current economic theories on union effects are based on capitalist markets in developed countries, they may not explain union effects in China, which is the world's largest developing country and is undergoing an economic system transition. Hence, empirical results from China can provide new insights that enrich economic theories concerning unions.

This study contributes significantly to the literature in two ways. First, although some studies have revealed that the union membership density and union wage premium differ between the public and private sectors (Blanchflower and Bryson 2010; Rosenfeld and Denice 2019), they have not investigated how differences in union membership density and union wage premiums affect the wage gap between the public and private sectors. This study is the first to explore the effects of two components of the wage gap simultaneously, based on the Blinder-Oaxaca (B-O) decomposition method (Blinder 1973; Oaxaca 1973). Second, regarding the estimations of the union wage premium, most studies have used single-time-point cross-sectional survey data and have not addressed the individual heterogeneity problem, except for some studies from developed countries (e.g., Freeman 1980). Based on national longitudinal survey data from five waves (2010, 2012, 2014, 2016, and 2018) of the China Family Panel Studies (CFPS), this study attempts to address this issue, using a fixed-effects (FE) model that has been used in empirical studies for developed countries (e.g., Freeman 1980). I also use a selection- bias-adjusted model with union lagged variables (LV), and propensity score matching (PSM) to address other endogeneity issues. Although this study does not tackle the endogeneity issue completely, it provides valuable empirical data, compared with previous studies, especially in the Chinese context.

The remainder of this paper is organized as follows. Section 2 introduces the institutional background and develops the two testable hypotheses. Section 3 introduces the methodology including the models, data, and variables used. Section 4 describes the results of the descriptive statistics. Section 5 reports and explains the results of the econometric analysis. Finally, Sect. 6 concludes the study.

2 Institutional background, theory, and hypothesis development

2.1 Institutional background

The function of unions in China changed with the SOE reform. During the 1949–1977 planned-economy period, the All-China Federation of Trade Unions (ACFTU), a national union organization, was formed and managed by the government. Subsequently, the Chinese government promulgated the first *Trade Union Law* in June 1950. The *Resolution on the Work of the All-China Federation of Trade Unions,* published on December 22, 1951, stated the following:

It is necessary to make all trade union workers further clearly understand the party's importance to unions, and each level of the ACFTU must work under the unified leadership of the committees of the Communist Party of China (CPC) at the same level…Under the people's democratic system, the most basic and important task of unions is to organize and educate workers to increase labor productivity, ensure the completion of the national production plan, and strive to overfulfill; on the basis of improving the production, it is necessary to frequently care for the daily needs of workers, provide services to fulfill these needs, and lead the workers in the struggle of socialism.

During the planned-economy period, the ACFTU established branches at the regional or industrial level, while regional unions managed industry-level unions (Guo and Dai 2022). Thus, all unions were managed by the ACFTU. However, unlike unions in Western countries, the ACFTU did not have the right to collectively bargain with employers for both employees' wage levels and employment because wage-setting and employment were managed by the central government during this period (Ma 2018).

Subsequently, the Chinese government implemented market-oriented reforms, beginning in 1978, and, since the late 1990s, has enforced the SOE reform (which promotes the privatization of middle- and small-sized SOEs, as well as enforces to establish modern corporate governance in SOEs). The government has given SOEs more autonomy for wage-setting and employment while simultaneously promoting the development of enterprises in the private sector. Since then, privately owned enterprises (POEs) and foreign investment enterprises (FIEs) have developed considerably (Lin et al. 1994, 2020).[1]

With the progress of market-oriented reforms, labor disputes about wage levels and employment have increased in China.[2] To address such new problems in the labor market, the Chinese government promulgated the Second *Trade Union Law* in April 1992, the *Labor Law of the People's Republic of China* in 1995, and the *Labor Contract Law of the People's Republic of China* in 2008. These regulations stated that "*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* on October 27, 2001, and the 32nd meeting of the Committee of the Thirteenth National People's Congress passed a revision of the *Trade Union Law* on December 24, 2021. Article 2 of the 2021 *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.*" Furthermore, Article 6 added the provision that "*protecting the legitimate rights and interests of employees is one of the basic responsibilities of the ACFTU.*" The enterprises eligible for these regulations include those in both the public and private sectors (You 2017; Guo and Dai 2022). Additionally, in contrast with developed countries (e.g., managers cannot become union members), according to China's *Trade Union Law*, both managers and non-manager employees can choose to join the firms or organizations' unions voluntarily.

During the current market-oriented reform period, according to the regulations mentioned above, although firm-and industrial-level trade unions are still managed by the ACFTU, they have obtained collective agreement rights and can influence wage-setting and employment at certain levels through collective agreements with employers. The negotiation usually takes place at the firm level. A given firm's union must report the collective agreement results to a higher-level ACFTU branch (e.g., regional branch of the ACFTU).

Furthermore, there has been a difference in union membership density between the public and private sectors during this period. As unions have been established in the public sector, including SOEs and public organizations (e.g., public school, public hospital, public research institute), and have been supported by the government[3] since the planned-economy period, most regular

[1] For example, the number of employees in non-SOEs (including POEs, FOEs, and self-employed individuals) increased from 86.67 million in 1990 to 522.13 million in 2019, while that of SOEs decreased from 103.46 million in 1990 to 82.61 million in 2019 (National Bureau of Statistics 2020).

[2] Based on the China Labor Statistics Yearbook (National Bureau of Statistics and Ministry of Labour and Social Security 1992–2022), the total number of laborers involved in labor disputes increased from 17,140 in 1992 to 1214,328 in 2008 and 1404,754 in 2021.

[3] For example, the union fund is composed of five parts: (1) union member fees from workers (approximately 10 CNY per month); (2) union fees from corporations or organizations (2% of the total employees' wage bills); (3) profits from SOEs and public organizations that are sent to the government; (4) government subsidies; and (5) miscellaneous items.

workers in this sector are likely to join unions. Although the ACFTU has enforced the establishment of union branches in POEs and FIEs since 2000, the union membership density in the private sector is still lower than that in the public sector (You 2017; Guo and Dai 2022).

2.2 Hypothesis development

In this section, I formulate two hypotheses based on economic theories and the institutional background mentioned for the empirical study above.

When considering the union wage premium and its disparity between the public and private sectors, two adverse effects can be seen, which will be explained below.

First, based on monopoly bargaining theory (Lewis 1963; Olson 1965), a union can protect union members (e.g., raising the wage level of union members) through collective bargaining with the employer; therefore, there is a positive union wage premium (e.g., Freeman 1980; Parsley 1980; Freeman and Medoff 1984; Hirsch and Addison 1986; Card 1996; Farber et al. 2021; Kulkarni and Hirsch 2021; Masso et al. 2022). Thus, higher union membership density is the tendency to increase the union wage premium.

If the government promotes the establishment of unions and expands union coverage in the public sector, this may increase unions' bargaining power significantly, which may lead to the union wage premium in the public sector becoming greater than that in the private sector. Blanchflower and Bryson (2010) and Rosenfeld and Denice (2019) report that the union density or collective agreement coverage rate in the public sector is higher than that in the private sector in the US and the UK, and the union wage premium in the public sector is also greater.

In China, union membership density in the public sector is greater than in the private sector, which is similar to the situation in developed countries. For example, based on data from the CFPS, the average proportion of union membership was 9.47% for the overall sample, 27.7% for the public sector, and only 4.4% for the private sector during the period 2010–2018. Therefore, it is predicted that the bargaining power of the public sector will be greater than that of the private sector, which may significantly increase the union wage premium (have an increasing effect).

Second, compared with developed countries, two reasons may decrease the union wage premium of China's public sector (have a decreasing effect). One reason is the difference in the target of a union's collective agreement between the public and private sectors. Specifically, in China, based on the regulations mentioned above, the ACFTU in the public sector aims to improve the work conditions for all employees, including both union members and non-members (You 2017; Guo and Dai 2022), whereas private sector unions target only union members, as is the case in developed countries. As the union's influence on non-members is greater in the public sector than in the private sector, the wage gap between union members and non-members may be narrower in the public sector than in the private sector.

The second reason is the government (or the Communist Party of China)'s influence on human resource management, including wage-setting, which is greater in the public sector than in the private sector (Lin et al.1994, 2020; Iwasaki et al. 2020; Ma and Iwasaki 2021; Jin et al. 2022). Consequently, the influence of unions' collective agreement on wage-setting may be smaller in the public sector, which means that the positive union wage premium may also be smaller in the public sector.

When the decreasing effect is greater than the increasing effect, the union wage premium in the public sector may be smaller than that in the private sector. Therefore, this study proposes hypothesis 1 (H1) as follows:

H1 In China, there is a positive union wage premium in both the public and private sectors. The premium in the private sector may be greater than that in the public sector.

Regarding the channels of unions' impact on the wage gap in the public and private sectors, this study estimates the effects of two channels: the difference in union membership density (the endowment effect) and the difference in the union wage premium (the price effect).

First, a union with greater membership density may have greater bargaining power, which may increase the probability of increasing union members' wage level (e.g., Lewis 1963; Freeman 1980; Farber et al. 2021). Blanchflower and Bryson (2010) and Rosenfeld and Denice (2019) report that the union wage premium in the public sector is much more significant than that in the private sector in the US and the UK due to that union membership density in the public sector is greater than that in the private sector.

Although the collective agreement right and power are smaller for Chinese unions than those in developed countries, the former also aim to improve working conditions, including raising the wage level. During the planned-economy period, in the public sector, the ACFTU was established, and the government enforced workers' participation in unions. Conversely, in the private sector, the establishment of unions has been permitted since the

1990s (You 2017; Guo and Dai 2022). Thus, the public sector has greater union membership density than the private sector, which may widen the wage gap between both sectors.

Second, the magnitude of the union wage premium may differ between the public and private sectors, which affects the wage gap between them. If the union wage premium in the public sector is greater than that in the private sector, it may widen the wage gap, and vice versa.

On the one hand, a higher union membership density may lead to greater bargaining power for the public sector, which may increase this sector's union wage premium and widen the wage gap between the public and private sectors (increase wage premium effect). Empirical studies have revealed that in developed countries, a higher union membership density in the public sector increases the union wage premium (Blanchflower and Bryson 2010; Rosenfeld and Denice 2019). It is anticipated that in China, a higher union membership density in the public sector may also increase the union wage premium in that sector. This, in turn, could contribute to the widening wage gap between the public and private sectors, similar to what occurs in developed countries.

On the other hand, in contrast with developed countries, the union wage premium may be smaller in China's public sector. This could be attributed to two reasons: First, the ACFTU aims to improve the work conditions for all employees, including union members and non-members (Yuan 2015; You 2017; Guo and Dai 2022). Second, the government's considerable influence on wage level in the public sector is greater than that in the private sector (Lin et al.1994, 2020; Iwasaki et al. 2020; Ma and Iwasaki 2021; Jin et al. 2022).

Considering that unions' main functions in the public sector include improving employees' working conditions and protecting their rights (for both union members and non-members), the decrease in the wage premium effect may be greater than the increase effect, which may decrease the union wage premium in the public sector. Therefore, generally, the difference in the union wage premium between the public and private sectors (where the union wage premium in the public sector is lower than that in the private sector) may conditionally narrow the wage gap.

Consequently, I propose hypothesis 2 (H2):

H2 Both union membership density and union wage premium differences conditionally affect the wage gap between the public and private sectors in China. The former widens the wage gap between the public and private sectors, while the latter narrows it.

I test H_1 and H_2 empirically in the following section.

3 Methodology

3.1 Models

3.1.1 Models for testing H1

This study uses a wage function to calculate the union wage premium in the total sample (Eq. (1)) and in the public and private sectors separately (Eq. (2)) to prove H1. I used the Ordinary Least Squares (OLS) method as the baseline estimations, which is expressed by Eqs. (1) and (2).

$$lnW_i = a + \beta_U U_i + \beta_{Pub} Pub_i + \beta_{nX} \sum_{1}^{n} X_i + \delta_t + u_i, \tag{1}$$

$$lnW_i^j = a^j + \beta_U^j U_i^j + \beta_{nX}^j \sum_{1}^{n} X_i^j + \delta_t^j + u_i^j, \tag{2}$$

where subscript i is an individual; t represents the time year (2010, 2012, 2014, 2016, or 2018); superscript j represents the sector (public or private). lnW represents the logarithm of the hourly wage. U is a union-membership dummy; Pub is a public-sector dummy; X represents other factors (e.g., education, years of work experience, gender, and occupation, and provincial regional dummies) that may affect wage levels; n represents the types of X. β_U represents the union wage premium; β_{pub} indicates the public sector wage premium; β_X is the coefficient of X; δ_t denotes the time year fixed effect; a is a constant; and u is an error term.

There are some concerns on econometric issues in the OLS method based on the cross-sectional data. This study also uses a set of methods to address these issues. First, regarding the omitted variable issue, several empirical studies have utilized family background variables such as parental education (e.g., father's and mother's education) as instrumental variables (IVs) in the wage function for education (Lemke and Rischall 2003; Mishra and Smyth 2013; Gong 2019). As parental education attainment might influence their adult children's wages through various channels, such as investment in their children's education or other unobservable cognitive and non-cognitive abilities, I added the family background variables (e.g., father's and mother's education) and re-ran the OLS method to check the robustness of the results.

Second, two sample selection biases may persist in Eq. (1). The first bias pertains to workers' choice to work in the public or private sector, while the second one relates to workers' choice to join a union. For instance,

workers with parents in the public sector are more likely to work in the public sector and have a higher probability of joining unions. This may lead to a situation in which workers in the public or private sectors and those with union membership are not randomly selected. I used a bivariate selectivity approach (Dubin and McFadden 1984) to address both types of selection bias. I established four dummy variables (*father's CPC membership, mother's CPC membership, father working in the public sector,* and *mother working in the public sector*) as identifications in the probability function of working in the public sector or joining a union in the first-stage estimations.[4] The variables associated with family background (such as parental education, employment status, or membership in the CPC) are commonly employed as IVs in empirical studies focusing on individuals' labor market outcomes (Lemke and Rischall 2003; Mishra and Smyth 2013; Gong 2019). In this study, I discovered that parental CPC membership or workplace ownership status (working in the public or private sector) significantly impacted the likelihood of an individual acquiring union membership or entry into the public sector (see Additional file 1: Table S1). However, these variables demonstrated insignificant influence on an individual's wage.[5] Consequently, I utilized these four variables as identifications in the selection bias adjusted models. Nevertheless, it is important to note that these four identification variables cannot be regarded as the sole complete exogenous variables used in the IV method.

Next, I calculated the correction terms for each sector as $\lambda_{pub} = \phi(\beta N)/\Phi(\beta N)$ for the public sector and $\lambda_{pri} = \phi(\beta N)/[1 - \Phi(\beta N)]$ for the private sector. Furthermore, I calculated the correction term for the choice to join a union as $\lambda_u = \phi(\gamma M)/\Phi(\gamma M)$, and added these correction terms to the wage function, as expressed in Eqs. (3) and (4).

$$lnW_i = a + \beta_U U_i + \beta_U U_i + \beta_{Pub} Pub_i + \beta_{nX} \sum_{1}^{n} X_i + \gamma_{pub} \lambda_{pubi} + \gamma_u \lambda_{ui} + \delta_t + u_{i,} \tag{3}$$

$$lnW_i^j = a^j + \beta_U^j U_i^j + \beta_{nX}^j \sum_{1}^{n} X_i^j + \gamma_{pub}^j \lambda_{pubi}^j \left(or \gamma_{pri}^j \lambda_{prii}^j \right) + \gamma_u^j \lambda_{ui}^j + \delta_t^j + u_i^j, \tag{4}$$

Third, there may be reverse causality issue. For example, union participation may be affected by a high union wage premium. The lagged variable of the union membership dummy was used to address the reverse causality issue. I used data from the survey wave dating back two waves prior to the survey in question ($t - 2$) as the lagged period in this study. For example, lagged union membership U_{t-2} (e.g., union membership in 2014) was used to investigate union membership's influence on wages in the current survey year (2018, as surveys are conducted every two years). The LV model based on the OLS method is expressed by Eqs. (5) and (6).

$$lnW_{it} = a + \beta_U U_{it-2} + \beta_{Pub} Pub_{it} + \beta_{nX} \sum_{1}^{n} X_{it} + \delta_t + u_{it,} \tag{5}$$

$$lnW_{it}^j = a^j + \beta_U^j U_{it-2}^j + \beta_{nX}^j \sum_{1}^{n} X_{it}^j + \delta_t^j + u_{it}^j, \tag{6}$$

Fourth, u_{it} in Eq. (1) includes individual-specific and time-invariant factors (ν_i) and idiosyncratic errors (ε_{it}). An individual heterogeneity issue (e.g., unobservable personality, ability, and preference) may arise in the estimated results if ν_i is maintained. This study utilized the longitudinal survey data and panel data estimation methods (e.g., the FE/RE model) to address this problem. The FE/RE model is expressed in Eqs. (7) and (8):

$$lnW_{it} = a + \beta_U U_{it} + \beta_{Pub} Pub_{it} + \beta_{nX} \sum_{1}^{n} X_{it} + \delta_t + \nu_i + \varepsilon_{it}, \tag{7}$$

$$lnW_{it}^j = a^j + \beta_U^j U_{it}^j + \beta_{nX}^j \sum_{1}^{n} X_{it}^j + \nu_i^j + \delta_t^j + \varepsilon_{it}^j, \tag{8}$$

where subscript t represents time year, and δ_t denotes the time fixed effect. Some studies have found that individual heterogeneity considerably affects the union wage premium in developed countries (e.g., Freeman 1980), whereas empirical studies for China (Yao and Zhong 2013; Li and Xu 2014; Yuan 2015; Gunderson et al. 2016; Li and Song 2017; Yang et al. 2018; Sun and Liu 2019)

[4] The results of the probability function of working in the public sector or joining a union are presented in the Additional file 1: Table S1.

[5] The results of wage functions using these identifications are available upon request.

have not addressed the heterogeneity problem. This study aims to address this issue in the Chinese context. However, it should be noted that although the FE/RE model can address the problem due to unobservable time-invariant variables of an individual, its effectiveness relies on the variation in each variable. When there's minimal variation in union membership during the analyzed period (from 2010 to 2018 in this study), the union effect may not be precisely estimated.[6] Furthermore, samples showing changes in union membership might possess specific characteristics that may not accurately represent the overall union effects across the entire sample. I only used the FE/RE model to conduct the robustness checks.

Fifth, some studies (e.g., Farber et al. 2021; Goerke and Huang 2022) have highlighted the existence of a sorting effect within union membership. Observable factors significantly influence the likelihood of joining a union, leading to non-random sample distributions between union members and non-members. For instance, research by Freeman (1980), Freeman and Medoff (1984), and Farber et al. (2021) demonstrate that union density is higher among less-educated, Black, and low-wage workers compared to their counterparts—the well-educated, White, and high-wage workers. Consequently, the groups of union members and non-members are not randomly selected. To mitigate this issue, this study also employed the PSM method to conduct the robustness checks.

The PSM method attempts to reduce bias due to observable variables in treatment effect estimates obtained by comparing the outcomes of the units that received treatment to those that did not (Rosenbaum and Rubin 1983). Using PSM brings the observed data closer to randomized experimental data by matching and resampling to minimize selectivity bias and counterfactual states in the sample composition. To apply this method to calculate the average disposition effects, one must (1) select appropriate observable variables for resampling, (2) run probit or logit regression to estimate the propensity score, (3) match the propensity score based on the selected observable variables (covariates), and finally, (4) calculate the average treated effect on the treated (ATT) based on the matched samples, as in Eq. (9).

$$\mathrm{ATT} = \mathrm{E}\Big(lnW_{it}^{j}(1) - lnW_{it}^{j}(0)|\mathrm{Union} = 1\Big), \tag{9}$$

where $lnW_{it}^{j}(1)$ expresses the logarithm of union members' wage in sector j and $lnW_{it}^{j}(0)$ represents that of the non-members with a similar endowment (e.g., education, years of work experience, and occupation) as the union members in sector j. This study reports the ATT results in the following section.

PSM assumes strongly ignorable treatment assignment, where treatment variable assignments and latent dependent variables are independent when conditioning observed covariates. Since covariates are summarized into one variable (i.e., the propensity score), this variable can be used even when the covariate values in two groups do not overlap (common support). Drawing from previous studies (Freeman 1980; Freeman and Medoff 1984; Farber et al. 2021, etc.) and specific characteristics of the Chinese labor market, such as segmentation by the public and private sectors, as well as by the population registration system (*hukou*) (Zhang et al. 2016; Cheng et al. 2020; Lin et al. 2020; Jin et al. 2022), this study uses several variables as observed covariates. These include demographic factors (years of schooling, years of work experience and experience squared value, health status, gender, ethnicity, *hukou* status, marital status, and CPC membership), work-related factors (occupation, industry sector, workplace ownership sector), family background (father's and mother's CPC membership, father's and mother's employment status in the public sector), province-level region, and survey year. However, it should be noted that the PSM method cannot address the problem due to unobservable variables of an individual.

Lastly, considering that the union wage premium may differ based on wage distribution, I also utilized unconditional quantile regression (QR) (Koenker and Baset 1978) and fixed effect QR (Machado and Santos Silva 2019) models. The QR model is expressed by Eq. (10).

$$\begin{aligned} &\min\theta\left|\ln W_i - \beta(\theta)H_i\right| + (1-\theta)\left|\ln W_i - \beta(\theta)H_i\right|, \\ &\rho_\theta \in (0,1) \end{aligned} \tag{10}$$

where θ represents the wage distribution percentile that should be between 0 and 1, H is the variable including the union dummy and other controlled variables including time-specific effect and individual fixed effects. The independent variables are the same as those in Eq. (1). $\rho_\theta(.)$ represents a check function—a loss function that retrieves the θ-th sample quantile. I performed estimations for the public sector, private sector, and the total sample.

Positive coefficients of the union dummies (β_U) or the ATT that are statistically significant in the public and private sectors in Eqs. (1)–(10) indicate that H1 is supported.

[6] The number of samples is 1048 for those staying as union member, 5666 for those staying as non-union member, 101 for those transferring from union member to non-union member (moving out), and 407 for those transferring from non-union to union member (moving in) among the total samples used in this study.

3.1.2 Models for testing H2

I used a wage decomposition method to test H2. First, I used a basic B-O decomposition method (Blinder 1973; Oaxaca 1973), which is represented by Eqs. (11) and (12).

$$\overline{lnW}_{pub} - \overline{lnW}_{pri} = \sum \beta_{pub}\left(\overline{H}_{pub} - \overline{H}_{pri}\right) + \sum \left(\beta_{pub} - \beta_{pri}\right)\overline{H}_{pri}, \quad (11)$$

$$\overline{lnW}_{pub} - \overline{lnW}_{pri} = \sum \beta_{\text{pri}}\left(\overline{H}_{pub} - \overline{H}_{pri}\right) + \sum \left(\beta_{\text{pub}} - \beta_{\text{pri}}\right)\overline{H}_{pub}, \quad (12)$$

where *pub* and *pri* denote the public and private sectors, respectively; $\overline{lnW}_{pub} - \overline{lnW}_{pri}$ is the wage gap between the two sectors; H represents the factors that affect wages, including union membership, demographic factors, work-related factors, provincial level regions, and survey year, which are similar to the variables used in Eq. (1); β is the coefficient of each factor. $\sum \beta_{\text{pub}}(\overline{H}_{pub} - \overline{H}_{pri})$ or $\sum \beta_{\text{pri}}(\overline{H}_{pub} - \overline{H}_{pri})$ expresses the differences in endowment effect (e.g., difference in union coverage), $\sum(\beta_{\text{pub}} - \beta_{\text{pri}})\overline{H}_{pri}$ or $\sum(\beta_{\text{pub}} - \beta_{\text{pri}})\overline{H}_{pub}$ denotes the differences in price effect (e.g., union wage premium).

I conducted a set of decomposition analyses as robustness checks. First, I used Oaxaca and Ransom's (1994) decomposition method (the O-R method) based on the wage function using the RE model to address the index number problem,[7] as expressed in Eq. (13).

$$\overline{lnW}_{pub} - \overline{lnW}_{pri} = \beta^*\left(\overline{H}_{pub} - \overline{H}_{pri}\right) + (\beta^* - \beta_{pri})\overline{H}_{pri} + (\beta_{pub} - \beta^*)\overline{H}_{pub} \quad (13)$$

where β^* is a sector-neutral coefficient estimated based on wage functions using the entire sample. $(\beta^* - \beta_{\text{pri}})\overline{H}_{pri}$ represents the gap caused by the too-low endowment return of the private sector (known as "private sector loss"), and $(\beta_{\text{pub}} - \beta^*)\overline{H}_{pub}$, represents the wage gap generated by the too-high endowment return of public sector (known as the "public sector gain"). The sum of these two decomposition values represents the wage gap resulting from the differences in the price effect.

Second, unlike in developed countries, where only non-manager employees can join unions, in China, both managers and non-manager employees can join unions. Since most union members are non-manager employees who often require more protection from unions than managers, I conducted the B-O decomposition using only non-manager employee samples (excluding manager samples) for robustness checks.

Third, while the fundamental Mincer-type wage function (Mincer 1958; 1974) traditionally uses education and years of work experience as indicators of human capital, several studies (e.g., Gustafsson and Li 2000; Mishra and Smyth 2013) have used the occupation and industry sector as broader human capital proxies. However, education may influence the choice of types of occupations or industrial sectors. For instance, highly educated workers are more likely to become managers or technicians, leading to the bad control variable problem in regression analysis. To address this issue and perform robustness checks, following Ma (2022), I also applied the B-O decomposition method only using the basic human capital variables (excluding occupation and industrial variables). The results of wage functions based on the OLS method are used in these decomposition analyses.[8]

When the contribution rate of the union in the endowment effect has a positive value and the contribution rate of the union in the price effect has a negative value in these decomposition results, it means that the union membership density widens the wage gap between the public and private sectors in China, while and union wage premium differences narrows the wage gap. Therefore, H2 is supported.

3.2 Data and variables

Longitudinal data from the CFPS, a national longitudinal survey project conducted by Peking University since 2010, were used (CFPS 2022) in this study. The baseline survey was officially launched in 2010 in 25 provinces,[9] municipalities, and autonomous regions, which represent 95 percent of the total population in China. 14,960 households were successfully interviewed. Within these households, 33,600 adults (individuals aged 18 years and above) and 8,990 youths (individuals aged 14 years and younger) were interviewed during the first wave. Despite the availability of six waves of CFPS data, I opted to use

[7] The index number problem arises because the B-O decomposition method employs different reference groups (for example, Eq. (11) and Eq. (12) are estimated separately), leading to variations in the results between Eq. (11) and Eq. (12).

[8] I also used the results of wage functions based on the FE and RE model in the B-O decomposition analysis. The majority of these results closely resembled those derived from the OLS method. Detailed results can be found in Additional file 1: Tables S3.

[9] The CFPS did not include Xinjiang Uygur Autonomous Region, Tibet Autonomous Region, Qinghai Province, Inner Mongolia Autonomous Region, Ningxia Hui Autonomous Region, and Hainan Province in mainland China and Hong Kong, Macau, Taiwan regions in 2010 baseline survey. The proportion of the population in these excluded regions is approximately 5 percent of the total population in China.

data from five waves—2010, 2012, 2014, 2016, and 2018—excluding the 2020 wave in this study.[10] This choice was made due to the COVID-19 pandemic's outbreak at the end of 2019 in China, which had a significant impact on employment and wages.

I used data from the CFPS for several reasons. First, the CFPS includes information on union membership status, wage, demographic factors (e.g., education, age, gender, ethnicity, health, CPC membership, and marital status), work-related factors (e.g., occupation, industrial sector, and employment in the public or private sector), and province-level regional information, all of which are essential for the empirical study. Second, the CFPS reveals national longitudinal data that can be used in the RE/FE model and LV method to address certain endogeneity issues.

The CFPS samples include both agricultural and nonagricultural workers and unemployed individuals. The sizes of the total samples were 33,600 (2010), 33,598 (2012), 37,147 (2014), 36,892 (2016), and 37,354 (2018) individuals. Nonagricultural employee samples were used in this study. As the *People's Republic of China Labor Law* prescribes 16 years as the minimum working age in China and considering the mandatory retirement age in the public sector,[11] I selected individuals aged 16–60 years for this study. After deleting data from individuals with missing values, the total sample included 14,084 respondents, including union members (2057), non-members (12,027), public sector workers (5,496), and private sector workers (8588). The public sector includes SOEs, collectively owned enterprises (COEs), and public organizations in this study.

The key dependent variable is the logarithm of the hourly wage. Based on the questions, "How much did you earn in the past 12 months?" and "How long did you work per week in the past year?" I obtained information on annual wages and weekly work hours. Based on the questionnaire items in the CFPS, wages were established based on overall earned income. I calculated hourly wages based on wages and work hours. To address the effect of inflation, I used the annual Consumer Price Index (CPI) published by the National Bureau of Statistics of China to adjust the wage levels each year, using the CPI in 2010 as a standard.

I considered the following control variables (See Appendix Table A1). First, based on economic theories (e.g., human capital theory, discrimination hypothesis) and previous studies, I used years of schooling, years of work experience and experience squared, and health status variables[12] as the indicators of human capital. I used gender (1 = male, 0 = female), ethnicity (1 = Han majority, 0 = ethnic minority), household registration [*hukou*] (1 = urban *hukou*, 0 = rural *hukou*), and marital status (1 = has a spouse, 0 = other) as demographic factors. The CPC membership dummy (1 = CPC member, 0 = non-CPC member) was used to control for the political background on wages (McLaughlin 2017; Ma and Iwasaki 2021).

Second, regarding work factors, five occupation dummies (manager, technician, clerk, operation worker, and other occupations) and five industrial sector dummies (manufacturing and construction, traffic and communication, retail and hotel, service, and other industry sectors) were also constructed. The public sector dummy variable was constructed as "1 if working in SOEs, COEs or public organizations, 0 if otherwise."

Third, considering that the regional disparities in labor demand and supply, minimum wage level, and local government policies may affect wages, I used the 25 province-level region dummies.

Finally, I constructed five-year dummy variables to control for the influence of the business cycle and changes in the macroeconomic environment by period.

4 Results of descriptive statistics

Table 1 summarizes the differences in the mean values of each factor between the public and private sectors. The results of the *t*-test indicate that differences remain between both groups. For example, the union membership density rate (i.e., the proportion of union members in this study) is significantly higher for the public sector (28.1%) than for the private sector (6.9%); the proportion of CPC members is higher for the public sector (30.6%) than the private sector (7.7%). Differences in occupation, industry sector distribution, region, and years are also significant. Thus, these variables should be controlled for in econometric analysis.

Figure 1 displays the kernel density of wage distribution among union members and non-members in the public and private sectors, respectively. The proportion of employees in high-wage areas is higher for union members than for non-members in both sectors. The

[10] I also employed the six waves, including the 2020 wave, to estimate the union wage premium, and found that the results were similar to those obtained using only five waves.

[11] The mandatory retirement age in the public sector is 50, 55, and 60 years for female workers, female cadres, and male workers and cadres, respectively. I also used samples of workers aged 16–50 years to perform a robustness check, and the results were similar to those reported in this study. The results are available upon request.

[12] Based on the question "How do you feel about your health status?" I constructed a dummy variable of health as follows: 1 if the respondents answered "good health" or "very good health", 0 if the respondents answered "normal health," "poor health," or "very poor health.".

Table 1 Descriptive statistics of variables

	Public (a)	Private (b)	Gap = (a)-(b)	*t*-test *p*-value
Log of hourly wage	2.503	2.205	0.298***	0.000
Union	28.1%	6.9%	0.212***	0.000
Year of schooling	13.116	10.166	2.950***	0.000
Year of work experience	21.114	21.243	− 0.129	0.504
Ethnicity (Han)	95.1%	96.1%	− 0.010***	0.003
Health	38.7%	42.7%	− 0.040***	0.000
Urban	90.6%	64.1%	0.265***	0.000
Married	77.5%	71.8%	0.056***	0.000
Party	30.6%	7.7%	0.229***	0.000
Occupation				
Manager	8.7%	7.3%	0.015***	0.000
Technician	30.4%	11.1%	0.192***	0.000
Clerk	23.1%	9.8%	0.133***	0.000
Operator	21.6%	39.1%	− 0.175***	0.000
Other occupation	16.2%	32.7%	− 0.165***	0.000
Industry				
Manufacturing and construction	21.3%	44.0%	− 0.227***	0.001
Traffic and communication	9.0%	7.6%	0.014***	0.000
Retail and hotel	4.7%	21.5%	− 0.168***	0.000
Service	39.4%	19.9%	0.195***	0.000
Other industry	25.6%	7.0%	0.186***	0.000
Region				
East	48.0%	59.1%	− 0.112***	0.000
Central	32.1%	26.9%	0.052***	0.000
West	19.9%	14.0%	0.060***	0.000
Year				
y2010	31.1%	34.6%	− 0.035***	0.000
y2012	18.5%	25.3%	− 0.067***	0.000
y2014	20.1%	12.2%	0.078***	0.000
y2016	10.2%	15.9%	− 0.057***	0.000
y2018	20.1%	12.0%	0.081***	0.000
No. of samples	5496		8588	

Source: Calculated using data from the CFPS of 2010, 2012, 2014, 2016 and 2018

***$p < 0.01$,

**$p < 0.05$,

*$p < 0.1$

mean value of the logarithm of the hourly wage is higher for union members than for non-members in both sectors. Finally, the wage levels of both union members and non-members are higher in the public sector than in the private sector, suggesting that there remains a wage gap between the public and private sectors among union and non-members.

Regarding the differences in union membership density, I calculated the proportion of union members among total employees in each group as the indicator of union membership density[13] in the public and private sectors, respectively. The results are presented in Fig. 2. From 2010 to 2018, the union membership density was 13.9%–40.9% for the public sector, and 2.8%–5.7% for the private sector, suggesting that union membership density in the public sector is greater than that in the private sector, leading to a large difference in union membership density between both sectors.

It is observed that the union membership density in the public sector increased significantly, while the change in the private sector was minimal from 2010 to 2018. This can be attributed to the following reasons: since 2010, the Chinese government has promoted the signing of labor contracts and the establishment of the collective wage negotiation system in the public sector, both of which have become the main functions of unions (Wang 2014; You 2017; Guo and Dai 2022). On July 26, 2010, during the Fourth Session of the 15th Session of the ACFTU, the ACFTU proposed the universal establishment of union organizations and a universal collective wage negotiation system in enterprises, in accordance with the law. The enhancement of unions' functions and the expansion of union coverage policies may have contributed to the increase in union membership density in China. Since the Chinese government primarily focused on institutional establishment and policy implementation in the public sector, the change in union membership density in the public sector was greater than that in the private sector from 2010 to 2018.

Although the results from these descriptive statistics indicate a wage gap between the public and private sectors, the wage level is higher for union members than for non-members, and there are differences in union membership density between both sectors, suggesting that unions may affect the wage gap between the public and private sectors. However, other variables (e.g., demography or work-related factors) that may affect wages have not been considered in these results. I will conduct an econometric analysis to control for these factors in the following section.

5 Empirical results

5.1 Results of testing hypothesis H1

5.1.1 Baseline results based on the OLS method

Table 2 presents the baseline results of the union wage premium based on the OLS method for the national

[13] Since there was no survey item asking about the presence of a union in the workplace, I only gathered information regarding whether the respondent was a union member.

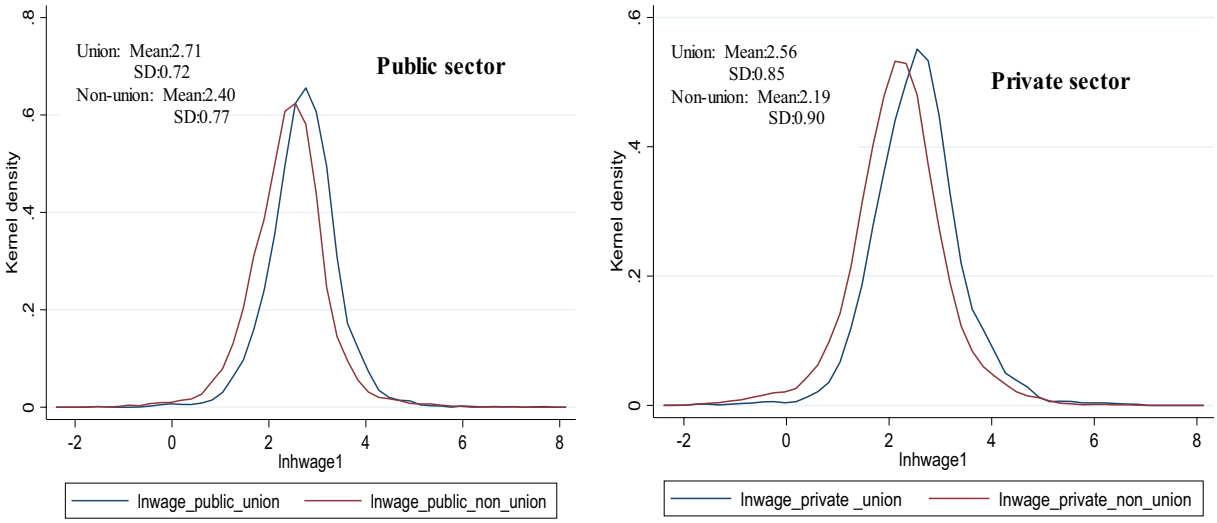

Fig. 1 Kernel density of wage distribution by union membership and by public and private sector. *Source*: Calculated using data from the CFPS of 2010, 2012, 2014, 2016 and 2018

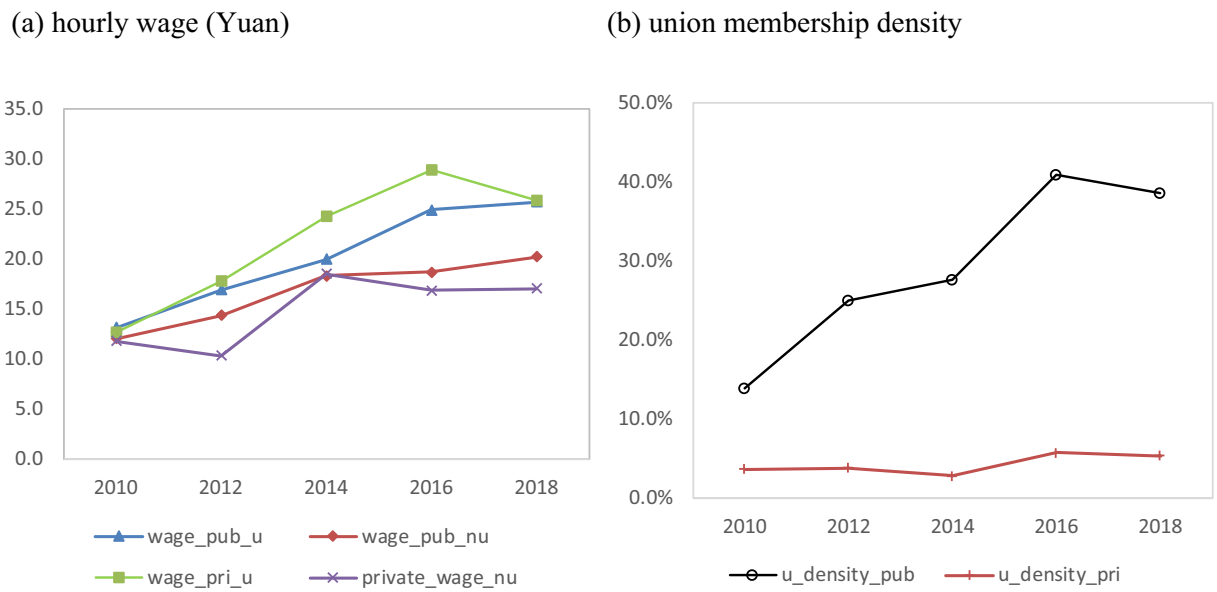

Fig. 2 Hourly wage and union membership density in public and private sector. *Source:* Calculated using data from the CFPS of 2010, 2012, 2014, 2016 and 2018. wage_pub_u: wage of union members in public sector; wage_pub_nu: wage of non-union members in public sector; wage_pri_u: wage of union members in private sector; wage_pri_nu: wage of non-union members in private sector; u_density_pub: proportion of union members in public sector; u_density_pri: proportion of union members in private sector

sample (Column 1), public sector (Column 2), and private sector (Column 3).[14]

First, regarding the union wage premium, the results based on the OLS method demonstrate that the union membership dummies are positive values and significant at 1% or 5% levels: 12.4% for the nation, 9.3% for the public sector, and 13.6% for the private sector. The results indicate that there is a significant positive union wage premium nationwide and in both the public and private sectors, while the premium of the private sector is greater than that of the public sector. Therefore, hypothesis H1 is supported.

[14] The detailed results of wage functions are presented in the Additional file 1: Table S2 (Model 1).

Table 2 Baseline results: Union wage premium in nation, public and private sector based on the OLS method

	(a) Nation	(b) Public	(c) Private
Union	0.124***	0.093***	0.136**
	(6.68)	(4.10)	(4.14)
Public	0.030*		
	(1.89)		
Control variables	Yes	Yes	Yes
No. of observations	14,084	5496	8588
R-squared	0.261	0.224	0.207

The *t*-values are in parentheses. The controlled variables include demographic factors (years of schooling, year of work experience, health status, sex, ethnicity, urban *hukou*, CPC member, married), work-related factors (occupation, industry sector, workplace ownership sector), province-level region and survey year (2012, 2014, 2016, and 2018) dummies, which are not presented in the table. These results are available upon request. *Source:* Calculated using data from the CFPS of 2010, 2012, 2014, 2016 and 2018.

***$p < 0.01$,

**$p < 0.05$,

*$p < 0.1$

To compare the results of this study with those of other studies pertaining to China, several scholars (Yao and Zhong 2013; Li and Xu 2014; Yuan 2015; Gunderson et al. 2016; Magda et al. 2016; Li and Song 2017; Yang et al. 2018; Sun and Liu 2019) have estimated the union wage premium in China using cross-sectional data. The results from these studies range from 5.5% (Yang et al. 2018) to 17.9% (Li and Xu 2014). The results of this study based on the OLS method are consistent with the findings of previous studies for China.

To compare the results of this study with those of developed countries, in contrast to the findings in Blanchflower and Bryson (2010) and Rosenfeld and Denice (2019) for the US and the UK, this study reveals that the union wage premium in the public sector is smaller than that in the private sector in China.

Two reasons may contribute to the results. First, the targets of a union's collective agreement differ between the public and private sectors. In China, the ACFTU in the public sector aims to improve the work conditions for all employees, including both union members and non-members (You 2017; Guo and Dai 2022), whereas private sector unions aim to protect only union members, as is the case in developed countries. In China, as the spillover effect of union on non-members in the public sector is greater than that in the private sector, the difference in the impact of union on wages between union members and non-members in the public sector is smaller than that in the private sector. Consequently, the union wage premium in the private sector is higher than that in the public sector in China. Second, the influence of the government on basic wage-setting in the public sector is greater than that in the private sector even in the current period (Iwasaki et al. 2020; Lin et al. 2020; Ma and Iwasaki 2021; Jin et al. 2022). Therefore, the influence of unions' collective agreement on wage-setting in the public sector may be smaller than that in the private sector, leading to a lower union wage premium in the public sector.

Second, concerning the wage gap between the public and private sectors, the coefficient of the public sector dummy is a positive value (0.030) and significant at a 1% level, suggesting the wage of the public sector is higher than that of the private sector even when the other factors (e.g., education, occupation, industry sector) are held constant. I will explore the determinants of the wage gap based on a decomposition method in the following section.

5.1.2 Sensitivity checks

I used three models based on the OLS method to perform sensitivity checks on the estimations of union wage premiums. The results are summarized in Table 3.

First, while union regulations in China allow both managers and non-manager employees to join unions, it is worth noting that unions in developed countries primarily focus on protecting non-managers rather than managers. Therefore, I restricted the sample to non-manager employees and conducted the estimations again (Model 1). The results reveal a significant positive union wage premium for the nation (8.3%), public sector (6.4%), and private sector (7.9%). The magnitude of the premium in the private sector is greater than that in the private sector. Consequently, hypothesis H1 was once again supported (Table 4).

Second, considering the influence of extreme values on wages, I adjusted the hourly wage (Model 2). I excluded samples with wage levels higher or less than three times the standard deviation and re-ran the estimations. The results indicate that there is a significant positive union wage premium for the nation (8.0%), and for the public (6.1%) and private sector (9.5%). The magnitude of the premium in the private sector is greater than that in the private sector. These results also supported hypothesis H1.

Lastly, I re-estimated the models by including the father's and mother's years of schooling. The results reveal a significant positive union wage premium for the nation (10.7%), public sector (8.43%), and private sector (10.7%). The magnitude of the premium in the private sector is greater than that in the public sector. These results further supported hypothesis H1.

Table 3 Sensitivity checks based on the OLS method

	(1) Nation	(2) Public	(3) Private
(1) Model1: using employee samples (non-managers)			
Union	0.083***	0.064***	0.079**
	(4.31)	(2.75)	(2.32)
Public	0.052***		
	(3.26)		
Control variables	Yes	Yes	Yes
No. of observations	12,992	5019	7973
R-squared	0.252	0.266	0.240
(2) Model2: using adjusted hourly wage			
Union	0.080***	0.061***	0.095***
	(5.05)	(3.02)	(3.56)
Public	0.096***		
	(7.24)		
Control variables	Yes	Yes	Yes
No. of observations	11,916	4488	7428
R-squared	0.189	0.176	0.185
(3) Model 3: adding family background variable			
Union	0.107***	0.843***	0.107***
	(5.46)	(3.45)	(3.10)
Public	0.016		
	(0.94)		
Father's education	0.007***	0.003	0.011***
	(3.87)	(1.04)	(4.25)
Mother's education	0.005**	0.004	0.006**
	(2.55)	(1.51)	(2.31)
Control variables	Yes	Yes	Yes
No. of observations	12,667	4997	7670
R-squared	0.219	0.219	0.214

The OLS method is used. The *t*-values are in parentheses. The control variables include demographic factors (years of schooling, year of work experience, health status, sex, ethnicity, urban *hukou*, CPC member, married), work-related factors (occupation, industry sector, workplace ownership sector), province-level region and survey year (2012, 2014, 2016, and 2018) dummies, which are not presented in the table. These results are available upon request. Model1: limited the samples to the employees (excluding managers); Model2: excluded the samples with the wage levels higher or less than three times standard deviations. *Source:* Calculated using data from the CFPS of 2010, 2012, 2014, 2016 and 2018.

***$p < 0.01$.

**$p < 0.05$.

*$p < 0.1$

5.1.3 Extending estimations using different models

There may exist estimated bias in the results based on the OLS method. I used several models to employ the robustness checks. First, in terms of sample selection bias (Model 1), the selection term for sector selection is significant for the private sector at a 5% level, and the selection terms for joining a union are significant at a 1% level for the nation, and the public and private sectors.[15] The results indicate that there are significant positive union wage premiums for the nation (5.1%) and the private sector (6.2%), while the premium is insignificant for the public sector. These conclusions align closely with those derived from the OLS method, reaffirming the support for hypothesis H1 once again.

Second, the results from the LV model (Model 2) based on the OLS method suggest.

that there are significant positive union wage premiums for the nation (8.8%) and the private sector (9.8%), while the premium is insignificant for the public sector. The conclusions of these results are also similar to those based on the OLS method. Thus, hypothesis H1 is confirmed again.

[15] The results of the probability functions based on the ransom-effects probit regression model are used to calculate the selection terms, which are presented in the Additional file 1: Table S1.

Table 4 Robustness checks using different models

	(1) Nation	(2) Public	(3) Private
(1) Model1: Selection bias adjusted model			
Union	0.051***	0.036	0.062***
	(2.75)	(1.15)	(2.73)
Public	0.117***		
	(5.25)		
Sector selection term	−0.059	−0.108	0.371**
	(−0.25)	(−0.28)	(2.44)
Union selection term	−0.860***	−0.646***	-1.366***
	(−9.83)	(−3.65)	(-6.46)
Control variables	Yes	Yes	Yes
No. of observations	14,084	5,496	8,588
R-squared	0.264	0.270	0.260
(2) Model2: LVt-2 model			
Union	0.088**	0.067	0.098*
	(2.52)	(1.51)	(1.71)
Public	0.013		
	(0.54)		
Control variables	Yes	Yes	Yes
No. of observations	3997	1801	2176
R-squared	0.255	0.214	0.308
(3) Model3: FE model			
Union	0.036	−0.026	0.086
	(1.22)	(−0.72)	(1.39)
Public	0.024		
	(0.63)		
Control variables	Yes	Yes	Yes
No. of observations	14,084	5496	8588
No. of groups	8242	2934	5701
R-squared within	0.158	0.200	0.131
Between	0.075	0.026	0.007
Overall	0.080	0.027	0.004

The *t*-values are in parentheses. FE: fixed effects model; LVt_2: lagged term (time t-2) of union dummy variable is used. The control variables include demographic factors (years of schooling, year of work experience, health status, sex, ethnicity, urban *hukou*, married), work factors (CPC member, occupation, industry sector), province-level region and survey year (2012, 2014, 2016, and 2018) dummies, which are not presented in the table. Parents' work status, parents' CPC membership dummies were used in first-stage estimation of sample selection bias adjusted model (Dubin and Mc Fadden, 1984). These results are available upon request. The level of standard error clustering is at the individual level in the FE model. *Source:* Calculated using data from the CFPS of 2010, 2012, 2014, 2016 and 2018.

***$p < 0.01$.

**$p < 0.05$.

*$p < 0.1$

Third, to address the individual heterogeneity issue, both the FE and RE models can be used. The results of the Hausman test[16] indicated that the FE model is more appropriate than the RE model. Therefore, I only report the results based on the FE model to conduct the robustness checks (Model 3).[17]

The results indicate that the coefficients of the union dummy variables are not statistically significant for the nation, and public and private sectors. This suggests that when addressing individual heterogeneity issues, there appears to be no significant union wage premium at a national level. Additionally, the disparity in union wage premiums between the public and private sectors is also found to be insignificant.

The results obtained from the FE model differ from those in Table 2 using the OLS method. Two reasons could explain these disparities. First, certain unobservable individual factors might significantly influence wages. When the FE model is employed to control for these time-invariant unobservable variables, the union wage premium might lose its significance. These unobservable variables encompass individual attributes such as innate abilities, personality traits, and individual preferences. Second, although the FE model can account for individual heterogeneity, its outcomes are contingent upon the changes in variables. When there is minimal variability in the variables over the analyzed period, the results may not accurately reflect the situations across the entire sample. For instance, the estimated coefficient of the union membership dummy in the FE model relies on the change of union membership during the study period (from 2010 to 2018 in this study). Given that only a small proportion of the sample experienced changes in union membership (7.37%), the results derived from the FE model might not comprehensively represent the overall union effect.

Furthermore, comparing the results based on the FE model between China and developed countries (e.g., Lewis 1963; Freeman 1980; Freeman and Medoff 1984; Card 1996; Farber et al. 2021; Kulkarni and Hirsch 2021; Masso et al. 2022) and Central Europe (Magda et al. 2016), while significant positive union wage premiums have been observed in developed countries (e.g., US, UK) and Central Europe, union wage premium is insignificant for China. The results of this study suggest that unions' impact on the wage-setting process to increase the wage levels of union members in China is smaller than that observed in other countries.

This may be attributed to differences in China and developed countries' protective targets and the functions of their unions. Specifically, in many developed countries, unions focus on providing benefits (e.g., increasing wage levels through collective agreements) to their union

[16] The results of Hausman test are 138.9, Prob > chi2 = 0.000 for public sector, 205.18, Prob > chi2 = 0.000 for private sector.

[17] The detailed results of wage functions are presented in the Additional file 1: Table S2 (Model 2).

Table 5 Results of union effect on wage for total, public sector and private sector based on the PSM method

	Sample	Treated	Controls	Difference	SE	t-value
(1) One to one Nearest-neighbor e matching (1:1)						
Nation	Unmatched	2.581	2.205	0.376	0.020	18.93
	ATT	2.580	2.463	0.117	0.031	3.78
Public	Unmatched	2.625	2.328	0.297	0.024	12.40
	ATT	2.624	2.548	0.077	0.037	2.10
Private	Unmatched	2.479	2.142	0.337	0.036	9.43
	ATT	2.479	2.357	0.122	0.052	2.33
(2) k-Nearest-neighbor matching (1:3)						
Nation	Unmatched	2.581	2.205	0.376	0.020	18.93
	ATT	2.580	2.460	0.120	0.026	4.62
Public	Unmatched	2.625	2.328	0.297	0.024	12.40
	ATT	2.624	2.530	0.094	0.031	3.04
Private	Unmatched	2.479	2.142	0.337	0.036	9.43
	ATT	2.479	2.373	0.106	0.044	2.40
(3) Kernel matching						
Nation	Unmatched	2.581	2.205	0.376	0.020	18.93
	ATT	2.581	2.483	0.098	0.023	4.20
Public	Unmatched	2.625	2.328	0.297	0.024	12.40
	ATT	2.625	2.534	0.091	0.028	3.21
Private	Unmatched	2.479	2.142	0.337	0.036	9.43
	ATT	2.479	2.359	0.120	0.038	3.13
(4) Radius matching						
Nation	Unmatched	2.581	2.205	0.376	0.020	18.93
	ATT	2.580	2.484	0.096	0.024	4.05
Public	Unmatched	2.625	2.328	0.297	0.024	12.40
	ATT	2.625	2.530	0.095	0.029	3.27
Private	Unmatched	2.479	2.142	0.337	0.036	9.43
	ATT	2.474	2.377	0.097	0.039	2.46
(5) Local linear regression matching						
Nation	Unmatched	2.581	2.205	0.376	0.020	18.93
	ATT	2.580	2.484	0.096	0.031	3.12
Public	Unmatched	2.625	2.328	0.297	0.024	12.40
	ATT	2.624	2.529	0.096	0.037	2.62
Private	Unmatched	2.479	2.142	0.337	0.036	9.43
	ATT	2.479	2.373	0.106	0.052	2.02

The observable variables include demographic factors (years of schooling, year of work experience, health status, sex, ethnicity, urban *hukou*, married, CPC membership), work factors (occupation, industry sector), parent's work status, parent' CPC membership, province-level region and survey year (2012, 2014, 2016, and 2018) dummies were used in the probit regression model to calculate the propensity matching score. Total samples including public and private sector. *Source:* Calculated using data from the CFPS of 2010, 2012, 2014, 2016 and 2018

PSM the propensity scores matching method. *ATT* average treated effect on the treated

***$p < 0.01$

**$p < 0.05$

*$p < 0.1$

members. While there may be a spillover effect benefiting non-members, the union wage premium among members typically exceeds that among non-members within corporations with a union. Conversely, Chinese unions, especially in the public sector, have broader targets that encompass both union and non-members. For instance, the *2021 Trade Union Law* stipulates that all branch organizations of the ACFTU should protect the legitimate rights and interests of overall employees, irrespective of union membership. Therefore, the union wage premium tends to be smaller in China. Additionally, in contrast to collective bargaining for wage increases

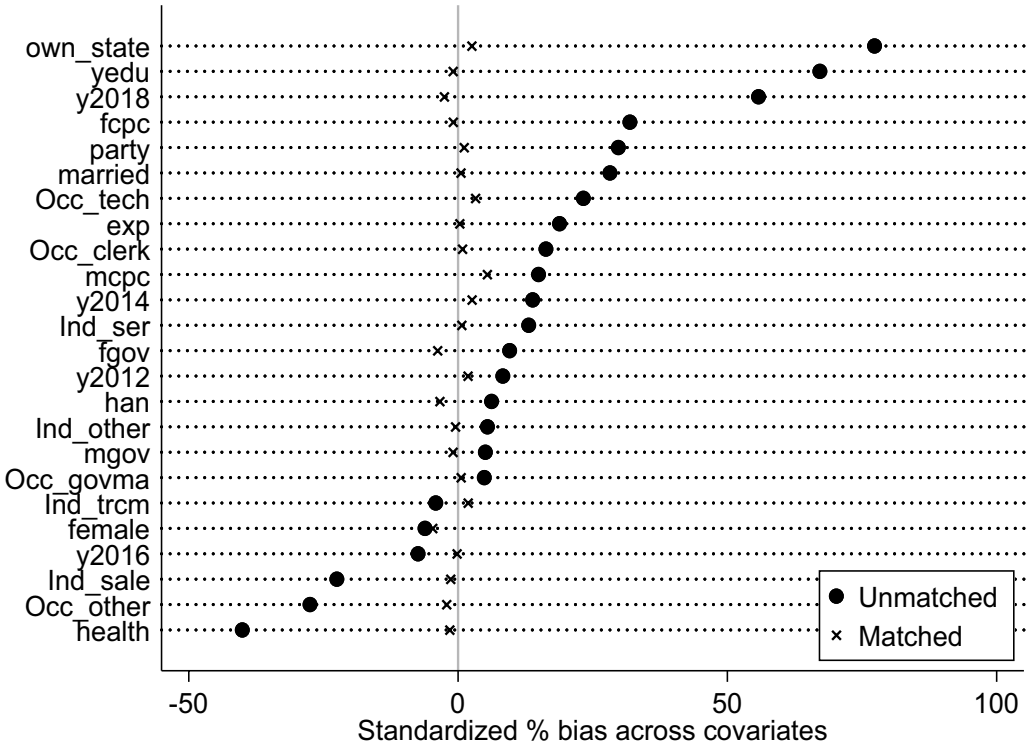

Fig. 3 Standardized bias across control variables among unmatched and matched group (total samples). *Source:* Calculated using data from the CFPS of 2010, 2012, 2014, 2016 and 2018. The results of province-level regional dummy variables were not presented in Fig. 3. These results are available upon request

with employers in developed countries, raising wage levels is not the primary function of Chinese unions. Article 1 of the *2021 Trade Union Law* describes the union's role as a bridge or link connecting the CPC organizations and Chinese workers. Given that the basic wage level in the public sector is controlled by the government, and the scope for increasing wages through collective agreements is limited in the private sector, the positive union wage premium is smaller in China.

Lastly, Table 5 summarizes the results obtained using several PSM methods. They are the one-to-one nearest neighbor matching (Model 1), *k*-nearest neighbor matching (1:3 matching in this study; Model 2), Kernel matching (Model 3), radius matching (Model 4), and local linear regression matching (Model 5) methods.

I estimated the probability function of joining unions for the overall sample as well as separately for the public and private sectors. These results were used to derive propensity scores.[18] The results (see Appendix Table A2) highlight significant influences of demographic factors, work-related factors, family background, region, and survey years on the probability of joining unions. For instance, the results indicate higher probabilities of joining unions among individuals in the public sector, Han majority workers, those with higher educational attainment, employees in manufacturing and construction sectors, and those whose fathers are CPC members. Moreover, there was a notable increase in the likelihood of joining unions in 2014, 2016, and 2018 compared to the reference year of 2010.

Unlike findings in advanced capitalist countries (Freeman and Medoff 1984; Farber et al. 2021), where union density is higher among disadvantaged groups—such as less-educated individuals, Black individuals, and low-wage workers—compared to advantaged workers like well-educated individuals and higher-wage earners, in China, the probability of union membership is higher among the advantaged group than among the disadvantaged group. This implies that union coverage among disadvantaged workers in China might be lower, suggesting that trade unions may not extend as many benefits

[18] Seven individual samples were dropped due to the resampling. The number of total samples used in the PSM method is 14,076.

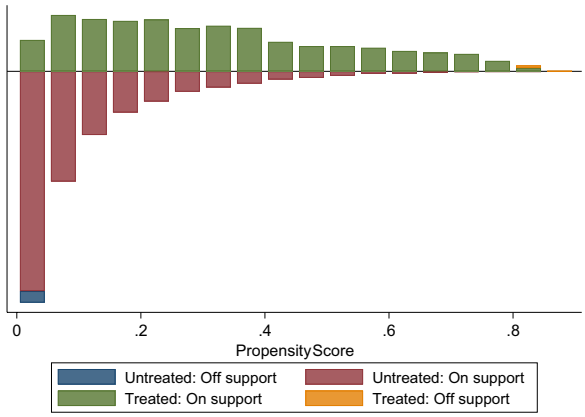

Fig. 4 Results of overlap checks (total samples). *Source:* Calculated using data from the CFPS of 2010, 2012, 2014, 2016 and 2018

to these workers, potentially contributing to a scenario where 'the rich get richer while the poor get poorer.

Figure 3 displays the standard bias across observable variables before and after matching for the balance check. The differences in the standard bias of each variable between union members and non-members became smaller (approaching a value near "0") after matching. Figure 4 presents the results of the overlap check. These results indicate that the matching procedures are appropriate.

The results from the five types of matching methods were similar. For example, the results based on one-to-one nearest neighbor matching (Model1) indicated that the result of ATT is 0.117, which is similar to that (0.124) in Table 2, suggesting there exists a positive union wage premium nationwide. The results support H1 again. Additionally, the values of differences in the matched group (0.117) became smaller than those in the unmatched group (0.376), suggesting that these observable factors may significantly affect the union wage premium through their effects on the union density.

To compare the results of the ATT between the public and private sectors, all results indicate that while the ATT values are significant positive values for both the public and private sectors, the magnitude of ATT values for the public sector is greater than that for the private sector. For instance, the results based on one-to-one nearest neighbor matching (Model1) indicated that the ATT value is 0.077 for the public sector and 0.122 for the private sector, indicating that while there exist positive union wage premiums for both the public and private sectors, the premium of the private sector is greater than that of the public sector. The results confirm hypothesis H1 again.

Table 6 Results of union wage premium by wage distribution percentile

	(1) QR			(2) QR_FE		
	Nation	**Public**	**Private**	**Nation**	**Public**	**Private**
5%	0.135**	0.156*	0.119	0.039	0.002	0.104
10%	0.086***	0.095**	0.106**	0.039	0.005	0.097
20%	0.067***	0.078***	0.057	0.038	0.010	0.095
30%	0.068***	0.071***	0.028	0.038	0.012	0.095
40%	0.076***	0.059***	0.048	0.037	0.016	0.094
50%	0.077***	0.056**	0.073**	0.036	0.025	0.075
60%	0.079***	0.037*	0.071**	0.034	0.036	0.059
70%	0.063***	0.042*	0.044	0.033	0.040	0.058
80%	0.088***	0.036	0.060	0.033	0.041	0.057
90%	0.091***	0.062	0.015	0.033	0.047	0.056
95%	0.068***	0.078	0.014	0.032	0.050	0.049

The controlled variables include demographic factors (years of schooling, year of work experience, health status, sex, ethnicity, urban *hukou*, CPC member, married), work-related factors (occupation, industry sector, workplace ownership sector), province-level region and survey year (2012, 2014, 2016, and 2018) dummies, which are not presented in the table. These results are available upon request. *Source:* Calculated using data from the CFPS of 2010, 2012, 2014, 2016 and 2018

QR quantile regression model; *FE_QR* fixed effects quantile regression model. 5 ~ 95% indicate the wage distribution percentiles (5% is the lowest wage level, 95% is the highest wage level)

***$p < 0.01$

**$p < 0.05$

*$p < 0.1$

Table 7 Basic decomposition results of wage gap between public and private sector

	(1)		(2)	
	Endowment (a)	**Price (b)**	**Endowment (c)**	**Price (d)**
Value: Total wage gap = 0.239	0.230	0.009	0.189	0.050
Contribution rate (%)				
Total	**96.0%**	**4.0%**	**78.7%**	**21.3%**
Each factor				
Union	**7.4%**	**− 1.3%**	**10.8%**	**− 4.7%**
Education	79.0%	139.8%	41.8%	177.0%
Experience	1.1%	93.3%	− 0.6%	94.9%
Health	− 0.2%	− 2.8%	− 0.6%	− 2.5%
Gender	1.6%	35.9%	4.4%	33.1%
Ethnicity	− 0.1%	3.3%	− 0.1%	3.2%
Married	− 1.9%	− 34.0%	0.5%	− 36.4%
Party	5.5%	1.9%	0.2%	7.1%
Occupation	8.6%	− 34.7%	28.8%	− 54.8%
Industry	− 1.1%	31.2%	− 4.9%	35.0%
Region	− 14.3%	0.9%	− 14.5%	1.1%
Year	10.4%	15.8%	12.8%	13.5%
Constant	0.0%	− 245.3%	0.0%	− 245.3%

Source: Calculated using data from the CFPS of 2010, 2012, 2014, 2016 and 2018. B-O decomposition method is used. Estimation 1 is based on Eq. (11), estimation 2 is based on Eq. (12) expressed in the text. The results of wage functions based on the OLS method are used in the decomposition analysis

5.1.4 Results by wage distribution

Considering unions' varying effects on the low-, middle-, and high-wage groups, I also calculated the union wage premium by wage distribution percentiles for both the public and private sectors using two models: (1) the QR model and (2) the fixed-effects QR model. The results are presented in Table 6.

The results obtained from the QR model indicate the presence of a positive union wage premium for the overall, public-sector, and private-sector samples, especially for the low- and middle-wage groups. Moreover, the magnitude of the union wage premium in the low-wage group in the public sector exceeded that in the private sector. These results support H1.

However, the results derived from the fixed-effects QR model suggest that, although the union wage premiums have positive values and these premiums are greater in the private sector across low-, middle-, and high-wage groups for the overall, public-sector, and private-sector samples, they are statistically insignificant. These results imply that when addressing the individual heterogeneity issue, the nation's union wage premium is insignificant.

5.2 Results of testing hypothesis H2

5.2.1 Basic decomposition results based on the B-O decomposition method

The basic decomposition results based on the B-O method are presented in Table 7. The results of wage functions based on the RE models were used in the decompositions.

First, the union effect on the endowment component was 7.4% (Column [a]) and 10.8% (Column [c]) respectively, implying that differences in union membership density contributed to widening the wage gap. In contrast, the union effect on the price component was − 1.3% (Column [b]) and − 4.7% (Column [d]) respectively, indicating that variations in the union wage premium reduced the wage gap. These results provide support for hypothesis H2.

Second, concerning the overall results of the endowment and price components, the contribution rate of the endowment component (96.0% in Column [a], 77.0% in Column [c]) exceeds that of the price component (4.0% in Column [b], 21.3% in Column [d]). This suggests that differences in endowment between the two sectors are the primary factors contributing to the wage gap.

Lastly, comparing the contribution rate of each factor, the years of education in the endowment component and the return to education and years of work experience in the price component are greater than those of the other factors. These results indicate that the influence of human capital is greater than that of other factors, which is consistent with the human capital theory (Becker 1964; Mincer 1974).

5.2.2 Robustness checks

In total, I conducted three robustness checks, and their results are presented in Table 8 and 9. First, I employed the O-R decomposition method. The results are displayed in Table 8. In Column [a], the union effect on the endowment component is 10.4%, suggesting that differences in union membership density widen the wage gap. In Column [d], the union effect on the price component is -4.3%, indicating that variations in the union wage premium narrow the wage gap. These findings suggest that both differences in union membership density and

Table 8 Robustness check (1): Decomposition using O-R decomposition methods

	Endowment		Price	
	(a)	(b) loss of non-union	(c) gain of union	(d) total (b + c)
Value: wage gap = 0.239	0.219	0.008	0.012	0.020
Contribution rate (%)				
Total	**91.4%**	**3.4%**	**5.2%**	**8.6%**
Each factor				
Union	**10.4%**	**− 0.2%**	**− 4.1%**	**− 4.3%**
Education	54.0%	45.9%	118.8%	164.7%
Experience	0.1%	35.5%	58.8%	94.3%
Health	− 0.4%	− 0.9%	− 1.7%	− 2.6%
Gender	3.4%	13.1%	21.1%	34.2%
Ethnicity	0.0%	− 0.4%	3.7%	3.3%
Married	− 0.5%	− 14.3%	− 21.1%	− 35.4%
Party	3.8%	1.2%	2.4%	3.6%
Occupation	23.1%	− 12.4%	− 36.7%	− 49.1%
I ndustry	0.0%	9.3%	20.8%	30.0%
Region	− 14.0%	1.4%	− 0.8%	0.7%
Year	11.7%	4.3%	10.2%	14.5%
Constant	0.0%	− 79.1%	− 166.3%	− 245.3%

Source: Calculated using data from the CFPS of 2010, 2012, 2014, 2016 and 2018. O-R decomposition method (Oaxaca-Ransom 1994) is used. The results of wage functions based on the OLS method are used in the decomposition analysis

Table 9 Robustness check (2): Decomposition using different samples and variables

	(1) Employees				(2) Excluding occupation and industry			
	Endowment (a)	Price (b)	Endowment (c)	Price (d)	Endowment (e)	Price (f)	Endowment (g)	Price (h)
Value: wage gap	0.233	0.023	0.204	0.052	0.243	0.079	0.236	0.086
Contribution rate (%)								
Total	**90.7%**	**9.3%**	**79.6%**	**20.4%**	**75.5%**	**24.5%**	**73.3%**	**26.7%**
Each factor								
Union	**6.1%**	**− 1.7%**	**11.0%**	**− 6.6%**	**5.7%**	**− 1.4%**	**10.0%**	**− 5.7%**
Education	69.0%	124.3%	34.9%	158.4%	70.1%	84.9%	45.5%	109.5%
Experience	0.4%	86.4%	− 0.1%	86.9%	0.3%	67.9%	0.1%	68.1%
Health	− 0.5%	− 1.2%	− 0.6%	− 1.1%	− 0.2%	− 0.6%	− 0.3%	− 0.6%
Gender	1.3%	34.8%	3.7%	32.4%	1.2%	31.9%	4.1%	29.1%
Ethnicity	0.0%	− 0.4%	0.0%	− 0.4%	− 0.1%	− 0.7%	− 0.1%	− 0.7%
Married	− 1.7%	− 31.5%	0.4%	− 33.6%	− 1.4%	− 31.7%	1.0%	− 34.1%
Party	5.6%	0.8%	3.1%	3.3%	2.7%	− 1.4%	6.7%	− 5.4%
Occupation	8.5%	-23.3%	28.3%	− 43.0%				
Industry	4.5%	21.6%	− 0.2%	26.3%				
Region	− 12.7%	1.1%	− 12.9%	1.3%	− 9.5%	4.9%	− 10.9%	6.4%
Year	10.2%	16.2%	12.2%	14.2%	6.6%	36.0%	17.2%	25.4%
Constant	0.0%	− 217.7%	0.0%	− 217.7%	0.0%	− 165.3%	0.0%	− 165.3%

Source: Calculated using data from the CFPS of 2010, 2012, 2014, 2016 and 2018. B-O decomposition method (Blinder 1973; Oaxaca 1973) was used. Estimations (a) and (c) were based on Eq. (11), estimations (b) and (d) are based on Eq. (12) expressed in the text. The results of wage functions based on the OLS method are used in the decomposition analysis

Table 10 Robustness check (3): Decomposition using three-fold B-O decomposition method

		Coef.	z
(1) Differential			
Prediction_1(Public)		2.408***	226.34
Prediction_2(Private)		2.169***	233.7
Difference		0.239***	16.96
(2) Decomposition	**Value**		**Percentage (%)**
Endowment	0.230***	15.87	96.2
Price	0.051***	2.85	21.3
Interaction	− 0.042***	-2.27	− 17.5

Source: Calculated using data from the CFPS of 2010, 2012, 2014, 2016 and 2018. Three-fold B-O decomposition method is used. The results of wage functions based on the OLS method are used in the decomposition analysis

differences in union wage premiums influence the formation of the wage gap between the public and private sectors. The former widens the wage gap, while the latter narrows it, thereby reaffirming hypothesis H2.

Second, I utilized the non-manager employee samples and conducted the B-O decomposition (Model 1 in Table 9). Columns [a] and [c] correspond to the results based on Eq. (11), while Columns [b] and [d] correspond to the results based on Eq. (12). The union effect on the endowment component is 6.1% (Columns [a]) and 11.0% (Columns [c]) respectively, indicating that differences in union membership density widen the wage gap. The union effect on the price component is − 1.7% (Columns [b]) and -6.6% (Columns [d]) respectively, suggesting that differences in union wage premiums narrow the wage gap. These results also support the notion that differences in union membership density contribute to widening the wage gap, while disparities in union wage premiums narrow the wage gap, thereby confirming hypothesis H2 once again.

Third, I excluded occupation and industrial variables from the controlled variables and conducted the B-O decomposition (Model 2 in Table 9). Columns [e] and [f] correspond to the results based on Eq. (11), while Columns [g] and [h] correspond to the results based on Eq. (12). The union effect on the endowment component is 5.7% (Columns [e]) and 10.0% (Columns [g]) respectively, suggesting that differences in union membership density widen the wage gap. The union effect on the price component is − 1.4% (Columns [f]) and -5.7% (Columns [h]) respectively, indicating that differences in union wage premiums narrow the wage gap. These results also provide further support for hypothesis H2.

Lastly, following Jann (2008), I utilized the three-fold decomposition method (Table 10). The endowment effect is calculated at 96.2%, while the price effect stands at 21.3%, and the interaction between both effects is − 17.5%. These results indicate that the impact of the endowment effect on the wage gap surpasses that of the price effect, which aligns with the findings in Table 7 (endowment effect: 85.2% and 77.0%; price effect: 14.8% and 23.0%). These results reaffirm hypothesis H2.

6 Conclusions

As trade unions are active in corporations worldwide, the union effects on the labor market have attracted global attention. However, empirical evidence is scarce regarding how trade unions' effects on the wage gaps differ between the public and private sectors, especially for China. Using national longitudinal survey data from five waves (2010, 2012, 2014, 2016, and 2018) of the CFPS and the Blinder-Oaxaca decomposition method, this study first estimates the effects of union membership density on the wage gap between the public and private sectors in China.

This study presents two key findings. First, the results based on the OLS method, selection bias adjusted model, LV model, and PSM method indicate the existence of a significant positive union wage premium nationwide, and the premium in the public sector is greater than that in the private sector. However, the premium becomes insignificant after addressing the issue of individual

heterogeneity. Second, the decomposition results indicate that the union membership density difference (the endowment effect) widens the wage gap between the public and private sectors; conversely, the union wage premium difference (the price effect) narrows the wage gap.

These findings have the following practical implications: the difference in union membership density widens the wage gap (Tables 7, 8 and 9). The proportion of union members in the private sector (2.8–5.7% in 2010 and 2018, respectively) was lower than that in the public sector (range: 13.9–40.9%) (Fig. 2). Therefore, a policy expanding union coverage in the private sector may effectively narrow the wage gap between both sectors. With the progress of market-oriented reforms, income inequality has expanded and become a severe issue (Sicular et al. 2021). Although Zhang and Cheng (2011) argued that because unions increase administration and bureaucratization processes, they cannot effectively improve China's income inequality problems, our results indicated that expanding the union membership density in the private sector can be expected to narrow the wage gap between the public and private sectors, which would contribute to reducing income inequality in China to a level comparable to that in developed countries (Card 1996; Farber et al. 2021).

Finally, this study had some limitations. First, while I examined union wage premiums in the public and private sectors and attempted to mitigate potential endogeneity issues by using the FE model as well as a selection-bias adjustment, LV and FE models, and the PSM method, it is important to note that the endogeneity issues could not be fully resolved. Further research is needed to explore the causal relationship between unions' effect on the wage gap between the public and private sectors. To address the individual heterogeneity issue, previous studies have predominantly utilized the FE model (e.g., Freeman 1980; Mellow 1981; Mincer 1981; Freeman and Medoff 1984; Walsh 2013; Ntlhola et al. 2019), a methodology also employed in this study. Apart from the FE model, several alternative econometric methods warrant consideration. Firstly, certain studies (Card 1996; Kulkarni and Hirsch 2021; Kölling 2022) utilize the information on the transformation of union membership following job changes. Due to the limited samples of changing union membership in the CFPS, I could not employ this method; however, future research could explore its application with more extensive observations from appropriate survey data. Secondly, in China, variations in union establishment between the public and private sectors, coupled with changes in unions in the private sector in response to evolving regulations and policies, suggest the potential application of quasi-experimental methods (e.g., difference-in-differences [DID], regression discontinuity design [RDD]). For instance, the DID method has been employed to investigate the causal effects of unions on wages in DiNardo and Lee (2004), Lee and Mas (2012), and Breda (2015) in developed countries.

Second, as Parsley (1980) and Booth and Bryan (2004) argued, a union has spillover and threat effects on nonmembers. These effects may differ between the public and private sectors and, hence, could influence the wage gap between them. Studying these effects presents new challenges for researchers.

Third, due to the limitations of the wage-related survey items in the 2010–2018 CFPS waves, this study solely estimated unions' effect on overall wage. Given that the composition of compensation components (e.g., basic wage, bonuses, and allowances) may vary between the public and private sectors, future research should aim to estimate the union wage effect by compensation type.

Fourth, while I have utilized variables such as occupation, industry sector, and ownership type from the CFPS to mitigate the influence of the workplace on workers' wages, it is noteworthy that other workplace factors like firm size, technological level, firm performance (e.g., debt, profit), and openness (such as tradable corporations) might impact workers' wages. The research to control these workplace factors based on employer-employee survey data has become a new issue (Breda 2015; Gürtzgen 2016; Masso et al. 2022). Furthermore, some demographic variables, such as personality traits and risk attributes, may affect the likelihood of choosing public or private sector and wage levels (Kamal and Blacklow 2022; Roethlisberger et al. 2023). The empirical study considering these factors based on appropriate survey data should be examined in future studies.

Despite these limitations, I believe that the current study, which took advantage of longitudinal data from the CFPS, provides insights into unions' effect on wages in the world's largest emerging market economy, while also offering new empirical evidence for understanding how unions affect the wage gap between the public and private sectors based on the decomposition methods.

Appendix

See Tables 11 and 12

Table 11 Definitions and descriptive statistics of variables

	Definitions	Mean	SD
Log of hourly wage	logarithm of the hourly wage calculated based on total wage and work hours	2.310	0.870
Union	1 = trade union member, 0 = non-member	0.095	0.293
Employment sector			
Public sector	1 = working in the public sector, 0 = working in the private sector	0.217	0.412
Education	Years of schooling	10.412	4.254
Experience	Years of experience = age-6-years of schooling	22.078	12.763
Ethnicity	1 = Han majority, 0 = ethnic minorities	0.943	0.232
Health	1 = health status is very good or good, 0 = otherwise	0.401	0.490
Urban	1 = urban *hukou*, 0 = rural *hukou*	0.420	0.494
Married	1 = having a spouse,0 = otherwise	0.771	0.420
Party	1 = a member of Communist Party of China,0 = non-CPC member	0.111	0.314
Occupation			
Manager	1 = manager or technician,0 = otherwise	0.095	0.294
Technician	1 = technician,0 = otherwise	0.131	0.337
Clerk	1 = clerk,0 = otherwise	0.098	0.298
Operator	1 = operator,0 = otherwise	0.383	0.486
Other occupation	1 = other occupation,0 = otherwise	0.293	0.455
Industry			
Manufacturing and construction	1 = working in the manufactural or construction industry sector,0 = otherwise	0.363	0.481
Traffic and communication	1 = working in the traffic or communication industry sector,0 = otherwise	0.066	0.249
Sales	1 = working in the traffic or sale industry sector,0 = otherwise	0.189	0.392
Service	1 = working in the traffic or service industry sector,0 = otherwise	0.220	0.414
Other industry	1 = other indusrty,0 = otherwise	0.162	0.368
Year			
y2010	1 = 2010 survey year, 0 = otherwise	0.145	0.352
y2012	1 = 2012 survey year, 0 = otherwise	0.154	0.361
y2014	1 = 2014 survey year, 0 = otherwise	0.223	0.416
y2016	1 = 2016 survey year, 0 = otherwise	0.225	0.417
y2018	1 = 2018 survey year, 0 = otherwise	0.254	0.435
Observations		14,084	

The samples were limited to those aged 16–60. Due to the limited space, the descriptive statistics of province-level region dummy variables were not presented in the table, they are available upon request. *Source:* Calculated using data from the CFPS of 2010, 2012, 2014, 2016 and 2018

Table 12 Results of probability function of joining union

	(1) Nation		(2) Public		(3) Private	
	Coef	z	Coef	z	Coef	z
Employment sector (Ref.: Private)						
Public	0.647***	18.96				
Education	0.115***	17.14	0.116***	12.22	0.117***	12.36
Exp	0.047***	7.98	0.047***	6.04	0.054***	5.98
Exp_sq	0.000***	− 3.68	0.000***	− 2.69	− 0.001***	− 3.13
Health	− 0.213***	− 6.63	− 0.221***	− 5.18	− 0.184***	− 3.83
Female	− 0.048	− 1.48	− 0.026	− 0.60	− 0.068	− 1.45
Han majority	0.194**	2.35	0.353***	3.35	− 0.044	− 0.35
Married	0.028	0.60	0.029	0.45	0.026	0.38
CPC membership	0.084**	2.11	0.138***	2.92	0.119*	1.68
Occupation (Ref.: Operator)						
Official and manager	− 0.169***	− 2.74	− 0.457***	− 5.37	0.105	1.23
Technician	− 0.141**	− 2.53	− 0.330***	− 4.49	0.090	1.04
Clerk	− 0.039	− 0.78	− 0.276***	− 4.21	0.298***	3.88
Other	− 0.162***	− 3.35	− 0.342***	− 4.93	0.004	0.06
Industry (Ref. Manu. and Cons.)						
Traffic and communication	− 0.173***	− 2.88	− 0.038	− 0.48	− 0.322***	− 3.35
Sale	− 0.133**	− 2.27	− 0.107	− 0.98	− 0.165**	− 2.33
Service	− 0.232***	− 4.95	− 0.124*	− 1.92	− 0.340***	− 4.93
Other	− 0.293***	− 6.15	− 0.189***	− 3.08	− 0.293***	− 3.70
Parent background						
Father: CPC member	0.170***	4.89	0.167***	3.75	0.161***	2.94
Mother: CPC member	0.086	1.24	0.053	0.62	0.142	1.26
Father in public sector	0.150	1.05	0.286*	1.73	− 0.159	− 0.49
Mother in public sector	0.164	0.55	0.052	0.15	0.451	0.83
Year (Ref.: y2010)						
y2012	0.291***	7.53	0.330***	6.39	0.224***	3.97
y2014	0.295***	6.76	0.362***	6.45	0.191***	2.79
y2016	0.317***	4.76	0.565***	5.37	0.091	1.03
y2018	0.785***	15.90	1.007***	15.70	0.400***	5.11
Province fixed effect	Yes		Yes		Yes	
Constants	− 3.647***	− 26.38	− 3.175***	− 16.58	− 3.534***	− 17.49
No. of sample	14,084		5,496		8,588	
Log likelihood	− 4641.980		− 2742.983		− 1977.211	
Pseudo R2	0.202		0.160		0.134	
Prob > chi2	0.000		0.145		0.000	

The probit regression model is used. *Source:* Calculated using data from the CFPS of 2010, 2012, 2014, 2016 and 2018.

***$p < 0.01$

**$p < 0.05$

*$p < 0$

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12651-024-00361-2.

Additional file 1: Table S1. Results of wage function based on the OLS method and FE model. **Table S2.** Results of probability functions used in selection bias adjusted method. **Table S3.** Decomposition results based on the B-O method and RE/FE model.

Acknowledgements
Not applicable.

Author contributions
XMA designed the study; collected the data; performed the formal analyses; and wrote, read, and approved the final manuscript.

Funding
This research was supported by the project grant of the Joint Usage and Research Center, Institute of Economic Research, Hitotsubashi University, Japan (Grant number: IERPK2323).

Data availability
The data used in this study can be obtained from https://opendata.pku.edu.cn/dataverse/CFPS?language=en. The datasets constructed by the author are available upon request.

Declarations

Ethics approval and consent to participate
The dataset used in this study, the China Family Panel Studies (CFPS), is publicly available (http://opendata.pku.edu.cn/en). The medical ethics committee approved the CFPS study, and all interviewees were required to sign informed consent. Ethics approval for the data collection in CFPS was obtained from the Biomedical Ethics Review Committee of Peking University (IRB00001052-14010). Ethics approval for the use of CFPS data was obtained from the University of Newcastle Human Research Ethics Committee. All the participants gave their written informed consent before any study procedures began.

Consent for publication
The author of this study declares that this manuscript does not infringe copyright, moral rights, or other intellectual property rights of any other person. The author of this study testifies that this paper is original and has not incurred any type of plagiarism. The author of this study declares that the article is original, has not already been published in a journal, and is not currently under consideration by another journal. The author of this study agrees to the terms of the Springer Open Copyright and License Agreement.

Competing interests
The author declares no conflicts of interest associated with this study.

Received: 20 June 2023 Accepted: 7 January 2024
Published online: 29 January 2024

References

Becker, G.S.: Human capital: A theoretical and empirical analysis, with special reference to education, 1st edn. Columbia University Press, Columbia (1964)
Blanchflower, D.G., Bryson, A.: The wage impact of trade unions in the UK public and private sectors. Economica **77**(305), 92–109 (2010)
Blinder, A.S.: Wage discrimination: Reduced form and structural estimation. J. Hum. Res. **8**, 436–455 (1973)
Booth, A., Bryan, M.: The union membership wage premium puzzle: is there a free rider problem? Ind. Labor Relat. Rev. **57**(3), 402–421 (2014)
Breda, T.: Firms' rents, workers' bargaining power and the union wage premium. Econ. J. **125**(589), 1616–1652 (2015)
Card, D.: The effect of unions on the structure of wages: a longitudinal analysis. Econometrica **64**(4), 957–979 (1996)
Cardozo, D.P., da Cunha, M.S.: Evidence of the public–private wage gap for Brazil in the period from 2008 to 2016. Ind. J. Labour Econ. **62**, 577–594 (2019)
Casale, D., Posel, D.: Unions and the gender wage gap in South Africa. J. Afr. Econ. **20**(1), 27–59 (2010)
Chang, J., Huang, H.: Trade unions, unemployment, economic growth, and income inequality. Macroecon. Dyn. **20**(1), 404–428 (2016)
Chen, G., Demurger, S., Fournier, M.: Wage differentials and ownership structure of China's enterprises. World Econ. Paper **6**, 11–31 (2005)
Cheng, H., Hu, D., Li, H.: Wage differential between rural migrant and urban workers in the People's Republic of China. Asian Dev. Rev. **37**(1), 43–60 (2020)
China Family Panel Studies (CFPS): https://opendata.pku.edu.cn/dataverse/CFPS?language=en. Accessed on May 20, 2022.
Choi, H., Ramos, R.: The union wage premium in a segmented labor market: new evidence from Korea. J. Asian Pac. Econ. **28**, 1979352 (2021)
Demurger, S., Li, S., Yang, J.: Earning differentials between the public and private sectors in China: exploring changes for urban local residents in the 2002s. China Econ. Rev. **23**, 138–153 (2012)
DiNardo, J., Lee, D.S.: Economic impacts of new unionization on private sector employers: 1984–2001. Q. J. Econ. **119**(4), 1383–1441 (2004)
Dubin, J., McFadden, D.: An econometric analysis of residential electric appliance holdings and consumption. Econometrica **52**(2), 345–362 (1984)
Farber, H.S., Herbst, D., Kuziemko, I., Naidu, S.: Unions and inequality over the twentieth century: new evidence from survey data. Q. J. Econ. **136**(3), 1325–1385 (2021)
Freeman, R.B.: The exit-voice tradeoff in the labor market: unionism, job tenure, quits and separations. Q. J. Econ. **94**(4), 643–673 (1980)
Freeman, R.B., Medoff, J. L.: What Do Unions Do? Basic Books (1984).
Gindling, T.H., Hasnain, Z., Newhouse, D., Shi, R.: Are public sector workers in developing countries overpaid? Evidence from a new global dataset. World Dev. **126**, 104737 (2020)
Gong, B.: Like father like son? Revisiting the role of parental education in estimating returns to education in China. Rev. Dev. Econ. **23**, 275–292 (2019)
Gunderson, M., Lee, B.Y., Wang, H.: Union pay premium in China: an individual-level analysis. Int. J. Manpow. **37**(4), 606–627 (2016)
Guo, H., Dai, W.: Original aspiration and mission: historical review and experience of Chinese trade unions participating in national governance. J. Beijing Federation Trade Unions Cadre Coll. **37**(4), 53–62 (2022)
Gürtzgen, N.: Estimating the wage premium of collective wage contracts: evidence from longitudinal linked employer–employee data. Indus. Relations **55**(2), 294–322 (2016)
Gustafsson, B., Li, S.: Economic transformation and the gender earnings gap in urban China. J. Popul. Econ. **13**(2), 305–329 (2000)
Hirsch, B. T., John T. A.: The Economic Analysis of Unions: New Approaches and Evidence. Allen and Unwin (1986).
Iwasaki, I., Ma, X., Mizobata, S.: Corporate ownership and managerial turnover in China and Eastern Europe: a comparative meta-analysis. J. Bus. Econ. **111**, 105928 (2020)
Iwasaki, I., Ma, X., Mizobata, S.: Ownership structure and firm performance in emerging markets: a comparative meta-analysis of East European EU member states, Russia, and China. Econ. Syst. **46**(2), 100945 (2022)
Jin, X., Xu, L., Xin, Y., Adhikari, A.: Political governance in China's state-owned enterprises. China J. Account. Res. **13**, 100236 (2022)
Kamal, M., Blacklow, P.: Attitudes and personality in the Australian gender wage gap. Appl. Econ. **54**(47), 5442–5459 (2022)
Kerr, A., Wittenberg, M.: Union wage premia and wage inequality in South Africa. Econ. Model. **97**(C), 255–271 (2021)
Koenker, R.W., Bassett, G.J.: Regression quantile. Econometrica **46**(1), 33–50 (1978)
Kölling, A.: Shortage of skilled labor, unions and the wage premium: a regression analysis with establishment panel data for Germany. J. Labor Res. **43**, 239–259 (2022)
Kulkarni, A., Hirsch, B.T.: Revisiting union wage and job loss effects using the displaced worker surveys. Ind. Labor Relat. Rev. **74**(4), 948–976 (2021)
Lee, D., Mas, A.: Long-run impacts of unions on firms: new evidence from financial markets, 1961–1999. Q. J. Econ. **127**(1), 333–378 (2012)
Lemke, R.J., Rischall, I.C.: Skill, parental income, and IV estimation of the returns to schooling. Appl. Econ. Lett. **10**(5), 281–286 (2003)

Lewis H. G.: Unionism and relative wages in the United States. The University of Chicago Press (1963).
Li, L., Song, Y.: Union wage effects on rural migrants: An empirical analysis by using propensity score matching. Chin. Rural Econ. **3**, 2–17 (2017)
Li, M., Xu, J.: Who benefits from China's trade union. Econ. Res. J. **5**, 49–61 (2014)
Lin, K.J., Lu, X., Zhang, J., Zheng, Y.: State-owned enterprises in China: a review of 40 years of research and practice. China J. Account. Res. **13**, 31–55 (2020)
Lin, Y., Cai, F., Li, Z.: Miracle Development Strategy and Economic Reform in China. Shanghai People Publishing (1994).
Ma, X.: Economic Transition and Labor Market Reform in China. Palgrave Macmillan (2018).
Ma, X.: Political background and its influences on wage gaps: Evidence from China. J. Asia Pac. Econ. 215712 (2022).
Ma, X., Iwasaki, I.: Does party membership bring a wage premium in China? A meta-analysis. J. Chinese Econ. Bus. Stud. **19**(1), 55–94 (2021)
Ma, X., Li, S.: Enterprise ownership reform and wage gaps between public and private sectors. In: Ma, X., Tang, C. (eds.) Growth mechanism and sustainable development of chinese economy: comparison with Japanese experiences. Palgrave Macmillan, London (2022)
Machado, J.A.F., Santos Silva, J.M.C.: Quantiles via moments. J. Econom. **213**(1), 145–173 (2019)
Magda, I., Marsden, D., Moriconi, S.: Lower coverage but stronger unions? Institutional changes and union wage premia in Central Europe. J. Comp. Econ. **44**, 638–656 (2016)
Maria, M., Christofides, L.N.: The impact of austerity measures on the public—private sector wage gap in Europe. Labour Econ. (2020). https://doi.org/10.1016/j.labeco.2020.101796
Masso, J., Meriküll, J., Vahter, P.: The role of firms in the gender wage gap. J. Comp. Econ. **50**, 454–473 (2022)
McLaughlin, J.S.: Does communist party membership pay? Estimation the economic returns to party membership in the labor market in China. J. Comp. Econ. **45**, 963–983 (2017)
Mellow, W.: Unionism and wages: a longitudinal analysis. Rev. Econ. Stat. **63**(1), 43–45 (1981)
Mincer, J.: Investment in human capital and personal income distribution. J. Pol. Econ. **66**(4), 281–302 (1958)
Mincer, J.: Schooling. Columbia University Press, Experience and Earning (1974)
Mincer, J.: Union effects: wages job turnover and job training. NBER Working Paper 808 (1981).
Mishra, V., Smyth, R.: Economic returns to schooling for China's Korean minority. J. Asian Econ. **24**, 89–102 (2013)
National Bureau of Statistics and Ministry of Labour and Social Security: China Labor Statistics Yearbook. Department of Population and Employment statistics National Bureau of Statistics, and Department of Planning and Finance, Ministry of Labour and Social Security (1992–2022).
National Bureau of Statistics of China: China Statistical Yearbook 2020. China Statistics Press (2020)
Ntlhola, M.A., Kwenda, P., Ntuli, M.: A distributional analysis of union-wage effects in South Africa: evidence from panel data. Dev. South. Afr. **36**(3), 309–328 (2019)
Oaxaca, R.L.: Male-female wage differentials in urban labor markets. Int. Econ. Rev. **14**(3), 693–709 (1973)
Oaxaca, R.L., Ransom, M.R.: On discrimination and the decomposition of wage differentials. J. Econom. **61**(1), 5–21 (1994)
Olson, M.: Logic of Collective Action: Public Goods and Theory of Groups. Harvard University Press (1965).
Parsley, C.J.: Labor union effects on wage gains: A survey of resent literature. J Econ Lit. **18**(1), 1–31 (1980)
Roethlisberger, C., Gassmann, F., Groot, W., Martorano, B.: The contribution of personality traits and social norms to the gender pay gap: a systematic literature review. J. Econ. Surv. **37**, 377–408 (2023)
Rosenbaum, P.R., Rubin, D.B.: The central role of the propensity score in observational studies for causal effects. Biometrika **70**(1), 41–55 (1983)
Rosenfeld, J., Denice, P.: What do government unions do? Public sector unions and nonunion wages, 1977–2015. Soc. Sci. Res. **78**, 41–56 (2019)
Shleifer, A., Vishny, R.W.: Politicians and firms. Q. J. Econ. **109**(4), 995–1025 (1994)
Shleifer, A., Vishny, R.W.: A survey of corporate governance. J Finance **52**(2), 737–783 (1997)
Sicular, T., Li, S., Yue, X., Sato, H. (Eds.) (2020) Changing Trends in China's Inequality: Evidence, Analysis, and Prospects. Oxford University Press (2020).
Sławinska, K.: Public–private sector wage gap in a group of European countries: an empirical perspective. Empir. Econ. **60**, 1747–1775 (2020)
Sun, Z., Liu, Y.: How does trade union influence workers' wage in China: an analysis based on CGSS 2008–2015 mixed cross-sectional data. Labor Econ. Stud. **7**(4), 121–144 (2019)
Walsh, F.: The union wage effect and ability bias: evidence from Ireland. Econ. Lett. **119**, 296–298 (2013)
Wang, Z.: In-depth thoughts on reforming and improving our country's income distribution system: discussing the role of trade unions in promoting collective wage negotiations. Shandong Trade Unions' Tribune **20**(1), 9–14 (2014)
Yang, Y., Mo, X., Tang, C.: The analysis of the "income premium" effect of labor union based on the national dynamic monitoring data of migrant population. Theory Pract. Finance Econ. **39**(5), 141–147 (2018)
Yao, Y., Zhong, N.: Unions and workers' welfare in Chinese firms. J. Labor Econ. **31**(3), 633–666 (2013)
Ye, L., Li, S., Luo, C.: Industrial monopoly, ownership and enterprises wage inequality: an empirical research based on the first national economic census of enterprises data. Manage. World **4**, 26–36 (2011)
You, Z.: On some quantitative studies into labor union in China. J. Southeast Univ. **19**(5), 60–66 (2017)
Yuan, Q.: The research of wage distribution under the union coverage effect and the union members effect: based on experience of employees employers matching data in 2012. J. Bus. Econ. **8**(286), 32–42 (2015)
Zhang, Y.: The change of income differentials between public and nonpublic sector in China. Econ. Res. **4**, 77–88 (2012)
Zhang, Y., Chen, J.: Effect of trade unions on industrial labor income in China. Asian Politics Policy **3**(1), 95–114 (2011)
Zhang, J., Xue, X.: State and non-state sector wage differentials and human capital contribution. Econ. Res. **4**, 15–25 (2008)
Zhang, L., Sharpe, R.V., Li, S., Darity, W.A., Jr.: Wage differentials between urban and rural-urban migrant workers in China. China Econ. Rev. **41**, 222–233 (2016)

Publisher's Note

Xinxin Ma Professor at the Faculty of Economics, Hosei University, Japan. Her current research work focuses on the labor market institutional transition and income inequality in China. The research papers have been published in, *China Economic Review, Economic Systems, Emerging Markets Review, Emerging Markets Finance and Trade,* and *Journal for Labor Market Research*. Dr. X. MA is the author of *Economic Transition and Labor Market Reform in China* (Palgrave Macmillan 2018), *Female Employment and Gender Gap in China* (Springer 2021).