

Zhao, Guochang

Working Paper

Narrow gender earnings gap in the urban areas of the People's Republic of China: To what extent does service economy matter?

ADB Working Paper Series, No. 1190

Provided in Cooperation with:

Asian Development Bank Institute (ADBI), Tokyo

Suggested Citation: Zhao, Guochang (2020) : Narrow gender earnings gap in the urban areas of the People's Republic of China: To what extent does service economy matter?, ADB Working Paper Series, No. 1190, Asian Development Bank Institute (ADBI), Tokyo

This Version is available at:

<https://hdl.handle.net/10419/238547>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by-nc-nd/3.0/igo/>



ADB Working Paper Series

**NARROW GENDER EARNINGS GAP
IN THE URBAN AREAS OF THE
PEOPLE'S REPUBLIC OF CHINA:
TO WHAT EXTENT DOES SERVICE
ECONOMY MATTER?**

Guochang Zhao

No. 1190
October 2020

Asian Development Bank Institute

Guochang Zhao is an associate professor of economics at the Research Institute of Economics and Management of Southwestern University of Finance and Economics, People's Republic of China.

The views expressed in this paper are the views of the author and do not necessarily reflect the views or policies of ADBI, ADB, its Board of Directors, or the governments they represent. ADBI does not guarantee the accuracy of the data included in this paper and accepts no responsibility for any consequences of their use. Terminology used may not necessarily be consistent with ADB official terms.

Working papers are subject to formal revision and correction before they are finalized and considered published.

The Working Paper series is a continuation of the formerly named Discussion Paper series; the numbering of the papers continued without interruption or change. ADBI's working papers reflect initial ideas on a topic and are posted online for discussion. Some working papers may develop into other forms of publication.

The Asian Development Bank refers to "China" as the People's Republic of China.

Suggested citation:

Zhao, G. 2020. Narrow Gender Earnings Gap in the Urban Areas of the People's Republic of China: To What Extent Does Service Economy Matter? ADBI Working Paper 1190. Tokyo: Asian Development Bank Institute. Available: <https://www.adb.org/publications/narrow-gender-earnings-gap-urban-areas-prc>

Please contact the authors for information about this paper.

Email: guochangzhao@swufe.edu.cn

Asian Development Bank Institute
Kasumigaseki Building, 8th Floor
3-2-5 Kasumigaseki, Chiyoda-ku
Tokyo 100-6008, Japan

Tel: +81-3-3593-5500
Fax: +81-3-3593-5571
URL: www.adbi.org
E-mail: info@adbi.org

© 2020 Asian Development Bank Institute

Abstract

It is believed that jobs in the service sector generally require more brain-related rather than labor-related skills compared to those in the industrial sector, thus females should have a comparative advantage in the service sector. Using the data from the 2017 China Household Finance Survey, I investigate the relationship between the service economy and the gender earnings gap in the urban People's Republic of China (PRC). The descriptive analyses show that the service economy hires more workers and pays more than the industrial sector on average; meanwhile the female employment share and gender earnings gap in the service sector are also higher and narrower, respectively, than those in the industrial sector. However, the decomposition analyses illustrate that the composition effect is only about 10% of the total gender gap, and that the service economy contributes most or all of the total composition effect across the earnings distribution. The service economy widens and narrows the gender gap through the composition effect and structure effect, respectively, but the latter is not insignificant. Therefore, simply increasing the share of the service economy is not likely to narrow the gender earnings gap in urban PRC. The fact that females are less likely to enter the high-pay service subsectors is the main reason why the rise of the service economy cannot narrow the gender earnings gap in urban PRC.

Keywords: gender earnings gap, service economy, OB decomposition, RIF regression, quantiles

JEL Classifications: J16, O14

Contents

1.	INTRODUCTION.....	1
2.	DATA, VARIABLES, AND DESCRIPTIVE ANALYSES.....	2
2.1	Data Sources and Sample Restriction	2
2.2	Variables and Descriptive Analyses	3
3.	METHODOLOGY.....	7
3.1	OB Decomposition.....	7
3.2	RIF Regression and RIF-OB Decomposition	8
3.3	OB and RIF-OB Decomposition with Reweighting.....	9
4.	RESULTS.....	11
4.1	OLS and Unconditional Quantile Regression Results	11
4.2	OB and RIF-OB Decomposition Results	16
5.	CONCLUSION AND DISCUSSION.....	19
	REFERENCES	20

1. INTRODUCTION

It is well-known that in the past three decades, the PRC has become the world's "factory," but a less noticeable fact is that the service economy has also developed quickly and the share of the service sector to the whole GDP had reached 54% up to 2019 (NBS of China 2020). The development of a service economy has a strong implication for the gender gap since compared to the industrial sector, the service sector needs fewer labor-related and more brain-related skills, and thus women have a comparative advantage in the service economy (Fan and Lui 2003; Weinberg 2000; Galor and Weil 1996; Goldin 2006; Rendall 2017; Ngai and Petrongolo 2017). In the same period, because of the education expansion, especially the higher education expansion since 1999, and the one-child policy since 1979, the female educational attainment in the younger generation has exceeded that of males on average (Tsui and Rich 2002; Lee 2012; Wu, Yan, and Zhang 2020). This equips women with the potential for working in the service economy, especially the high-pay service economic industries. This paper aims to investigate the relationship between the service economy and the gender earnings gap in the PRC's urban labor market.

The gender earnings gap is a classical research issue in labor economics. Previous studies more or less concentrated on the productive characteristics, including education (e.g., Blau and Kahn 2017; Liu 2011), cognitive and noncognitive abilities (e.g., Pope and Sydnor 2010; Ceci et al. 2014; Fortin 2008; Kuhn and Weinberger 2005; Borghans, ter Weel, and Weinberg 2014), and discrimination (Li, Song and Liu 2014). With regard to the gender earnings gap in the PRC, the reform of labor market institutions is also one of the most important factors influencing the gender earnings gap (Song, Sicular, and Gustafsson 2019). In recent years, due to the rise of the service economy and also the changes in technology, females have gained a comparative advantage in the labor market, especially in certain jobs in the service sector (Borghans, ter Weel, and Weinberg 2014). Based on these facts, Fan and Lui (2003), Ngai and Petrongolo (2017), and a few others began to analyze the gender earnings gap and its change from the viewpoint of industry composition.

Based on the data from the 2017 China Household Finance Survey, this paper investigates the service economy and gender earnings gap in urban PRC. I decompose the gender earnings gap using the classical Oaxaca-Blinder (OB) decomposition (Oaxaca 1973; Blinder 1973) and the OB decomposition based on recentered influence function (RIF) regression (RIF-OB) (Firpo, Fortin, and Lemieux 2007, 2009, 2018). The results show that the gender earnings gap is 0.23 log points on average, and it ranges from 0.10 to 0.33 log points over the earnings distribution. Compared to the industrial sector, the service economy hires more workers, has a higher female employment share, a lower gender earnings gap, and higher average annual earnings. However, only a small part of the total gender earnings gap is attributed to the composition effect, which is totally or mostly attributed to the service economy. The service economy widens the gender earnings gap through the composition effect, but narrows the gender earnings gap through the structure effect, though the latter is not statistically significant.

I then discuss the policy implications of these findings. The findings in this paper suggest that simply increasing the share of the service economy may widen the gender earnings gap, because currently females are more likely to work in the low-pay service economic subsectors. On the one hand, removing the barriers to females entering high-pay industry should be helpful in narrowing the gender earnings gap. On the other hand, females' decisions in the labor market and marriage market may be jointly made, and

thus it may be meaningful to consider the gender consumption and leisure differentials in addition to the gender earnings gap in future research.

My paper makes the following contributions. First, the gender earnings gap has widened since the 1980s, and it began to narrow from 2007 onwards (Liu 2011; Li and Song 2013; Chi and Li 2014; Song, Sicular, and Gustafsson 2019), and this paper updates the understanding of the gender earnings gap in urban PRC. Second, this paper contributes to the literature on service economy and the gender earnings gap. It is believed that the rise of a service economy should be able to narrow the gender earnings gap, and this idea obtained empirical support from studies on developed economies (for example, Fan and Lui 2003; Ngai and Petrongolo 2017), but this paper points out that this is not necessarily true. Compared to the industrial sector, the service sector employs more female workers generally but the earnings among different service economic subsectors are quite dispersed. If females enter the high-pay service subsectors with more difficulty than males, the expansion of the service economy may further widen rather than narrow the gender earnings gap.

The rest of the paper is constructed as follows. Section 2 describes data and variables and presents descriptive analyses. Section 3 introduces the methodology. Section 4 presents the empirical results and Section 5 provides some discussions about the policy implications of the main findings. Finally, Section 6 concludes.

2. DATA, VARIABLES, AND DESCRIPTIVE ANALYSES

2.1 Data Sources and Sample Restriction

This paper uses the data from the China Household Finance Survey (CHFS). The CHFS is a nationally representative longitudinal survey conducted by the Southwestern University of Finance and Economics, PRC. The CHFS began in 2011 and has since been conducted every two years. After the second wave, the sample size was expanded dramatically to obtain representativeness in 29 provinces/municipalities of the PRC.¹ This study uses the 2017 CHFS data, which were the most recently released when I began this study.

The CHFS 2017 data include 40,011 households, and 127,012 individuals. Given that we are investigating the relationship between the gender earnings gap and a service economy, and the service economy is mainly an urban economic activity, I first drop the rural observations, which excludes 45,067 observations. Second, I exclude individuals aged below 20 because most of these young people in urban areas are still in school, and exclude older people aged above 50 because the female retirement age is 50 and it is not meaningful to include the elderly beyond 50 when we study the gender earnings gap. This step drops a further 47,681 observations. Among the remaining observations, 26,511 people had a job during the past year, namely 2016, and 14,126 people were employees, who are the individuals of interest in this study.² Next, I exclude people with missing values on education, gender, age, occupation, and industry information, and exclude 50 individuals with annual earnings lower than

¹ Xinjiang and Tibet Autonomous Region are not included in the CHFS.

² One may wonder why I do not extend to the gender income gap rather than just the gender earnings gap. The main reason is that the CHFS data do not provide the individual income when an individual engages in a self-employment job or his/her household operate a business or private firm. Examining the gender income gap among these people is important, but the definition of individual income is not clear and it needs careful and tedious work to define and calculate the individual income for this subpopulation. Nevertheless, this should be one direction for extending this research in the future.

100 yuan. Finally, I obtain a sample of 12,915 observations, consisting of 6,812 males and 6,103 females.

2.2 Variables and Descriptive Analyses

The CHFS data contain three after-tax sources of earnings for employees: salary, bonus, and subsidies including cash and allowances in kind during the past year (2016). I add up these three sources and calculate the sum annual earnings, which is the dependent variable in this study.

The CHFS data adopt the Standard China Industry Classification System (GB/T4754), which groups all economic activities into 20 industries, and an additional group if it cannot be classified easily into any industry. Of these, Industry 1 is from the primary sector (the first industry), Industries 2–5 are from the industrial sector (the second industry), and the rest are from the service sector (the third or tertiary industry). Although the CHFS has a large sample size, the number of observations in any particular industry may still be small, and it might be misleading to ignore the substantial difference with the service sector and just view the service sector as a whole. Therefore, I reclassify the service sector into seven subsectors according to the female employment share in each industry. These seven service economy subsectors are: (1) transportation, logistics, and warehousing; (2) logical, mathematics, and/or engineering skill-intensive services – scientific research, technology services, geological prospecting, information, computer and software services, hydraulics, environment and public facility administration; (3) wholesale, retail, accommodation, and catering services; (4) finance and real estate; (5) education, medical, social security, and welfare; (6) public administration, social organization, and international organization; and (7) others.

The educational attainment is measured by a categorical variable in the CHFS data. Based on the education distribution and the education system in the PRC, I regroup the education system into four categories: junior high school or below, senior high school, three-year college and university or postgraduate.³ In the PRC, the household registration system classifies individuals into agricultural and nonagricultural *Hukous*, and in recent years there have been some other types such as the rural-urban unified Hukou. Marriage status in the CHFS comprises seven types, including single, first married, divorced, and so on. Based on the distribution among marriage status, I just define a dummy variable for single. The CHFS data classify the occupation of employees into six groups: management, profession, clerical staff, service staff, manufacturing workers, and other occupations.

Table 1 presents summary statistics by gender. The average annual earnings of males and females are approximately 65,000 and 51,000 yuan, respectively. In other words, males earned about 28% more than females on average in 2016. Females are 0.25 years younger than males. As for educational attainment, females slightly outperform males, which is consistent with previous findings (e.g., Li and Zhao 2020). Male workers are a little more likely to have a nonagricultural Hukou and be single than female workers.

³ In our sample, less than 4% of people have only a primary school education. Given the small share, it is combined with junior high school. The group of senior high school education also includes vocational education.

Table 1: Summary Statistics by Gender

Variables	Male		Female		Diff.
	Mean (1)	S.D. (2)	Mean (3)	S.D. (4)	(1)–(3) (5)
Annual earnings	65,071	69,059	50,830	53,261	14,241***
Industries					
Industrial sector	0.31	0.46	0.17	0.38	0.14***
Service sector	0.69	0.46	0.83	0.38	–0.14***
Service subsector 1	0.09	0.29	0.04	0.19	0.06***
Service subsector 2	0.09	0.29	0.05	0.22	0.04***
Service subsector 3	0.14	0.35	0.23	0.42	–0.09***
Service subsector 4	0.07	0.26	0.08	0.28	–0.01***
Service subsector 5	0.13	0.33	0.29	0.45	–0.16***
Service subsector 6	0.10	0.31	0.08	0.28	0.02***
Service subsector 7	0.05	0.23	0.05	0.22	0.00
Age	36.13	8.10	35.88	8.01	0.25**
Educational attainment					
Junior high school or below	0.18	0.39	0.19	0.39	0.01
Senior high school	0.23	0.42	0.21	0.41	–0.02***
Three-year college	0.23	0.42	0.22	0.42	–0.01
University or postgraduate	0.35	0.48	0.38	0.49	0.03***
Hukou					
Agricultural	0.26	0.44	0.26	0.44	0.01
Nonagricultural	0.60	0.49	0.58	0.49	–0.02**
Others	0.15	0.35	0.16	0.37	0.01**
Single	0.20	0.40	0.17	0.38	–0.03***
Occupations					
Management	0.06	0.24	0.04	0.19	0.02***
Professional	0.35	0.48	0.30	0.46	0.05***
Clerical staff	0.31	0.46	0.36	0.48	–0.06***
Service staff	0.12	0.33	0.17	0.37	–0.05***
Manufacturing workers	0.11	0.32	0.08	0.27	0.04***
Other occupations	0.05	0.23	0.06	0.23	0.00
N	6,812		6,103		

Note: The sampling weight is used when calculating the summary statistics. The service sector is classified into seven subsectors: (1) transportation, logistics, and warehousing; (2) logical, mathematics, and/or engineering skill-intensive services – scientific research, technology services, geological prospecting, information, computer and software services, hydraulics, environment and public facility administration; (3) wholesale, retail, accommodation and catering service; (4) financing and real estate; (5) education, medical, social security and welfare; (6) public administration, social organization, and international organization; (7) others. Significance codes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

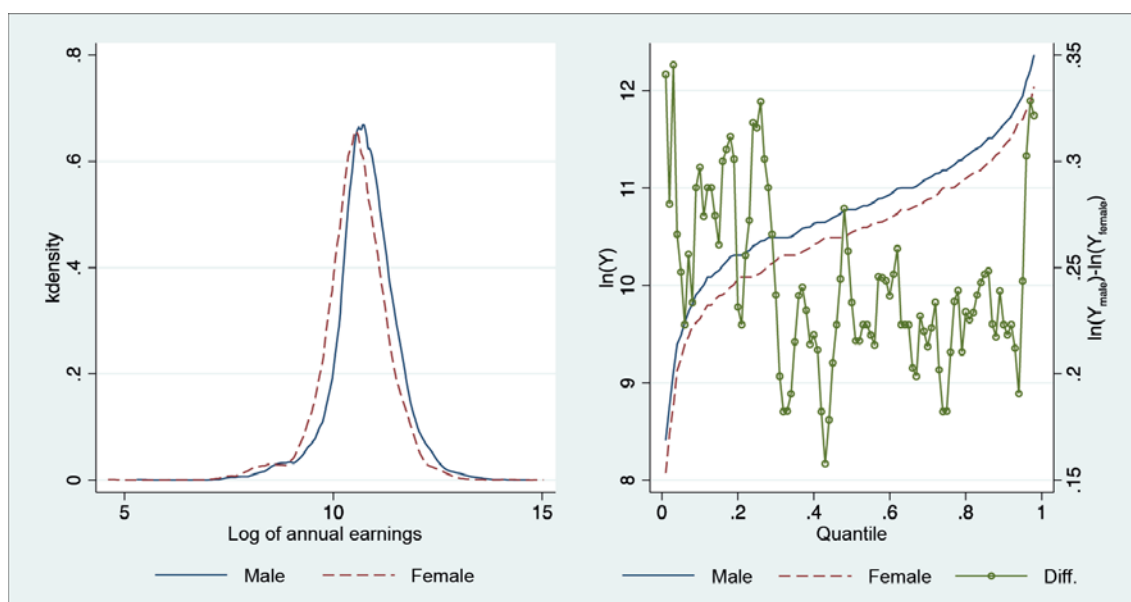
About 30% of males work in the industrial sector, while the corresponding ratio is 17% for females, which is consistent with our expectation since the jobs in industrial sectors are more likely to be brawn-intensive. Among the service sectors, subsector 3 (wholesale, retail, accommodation and catering services) and subsector 5 (education, medical, social security and welfare) employ most of the workers. This is the same for males and females, but female workers are concentrated more in these two subsectors. In addition, compared to females, male workers are also a little more likely to work in service subsector 1 (transportation, logistics, and warehousing) and subsector 2 (logical, mathematical and/or engineering skill-intensive services). The work in service subsector 1 is more likely to be brawn-intensive, as in the industrial sector, while for

service subsector 2 there is a stereotypical image that women are less capable than men in logical, mathematical, and engineering thinking, which may create a barrier to women engaging in these jobs in these industries.

In terms of occupation choices, females are a little less likely to engage in the occupations of management, professional and manufacturing workers, but a little more likely to be clerical and service staff. Nevertheless, the sizes of the differences are not as large as in industry.

In order to further investigate the gender earnings gap, I plot the distribution of logged earnings by gender and the gap at different quantiles in Figure 1. The graphs show that the gender earnings gap is larger at the lower quantiles and at the top end. From the bottom end to the 0.30th quantile, the gender earnings gap is about 0.3 log points, as it is at the top end. In contrast, the gap is about 0.22 log points from the 0.30th to the 0.92nd quantile. This means that a “sticky floor” and a “glass ceiling” both exist in the Chinese urban labor market.

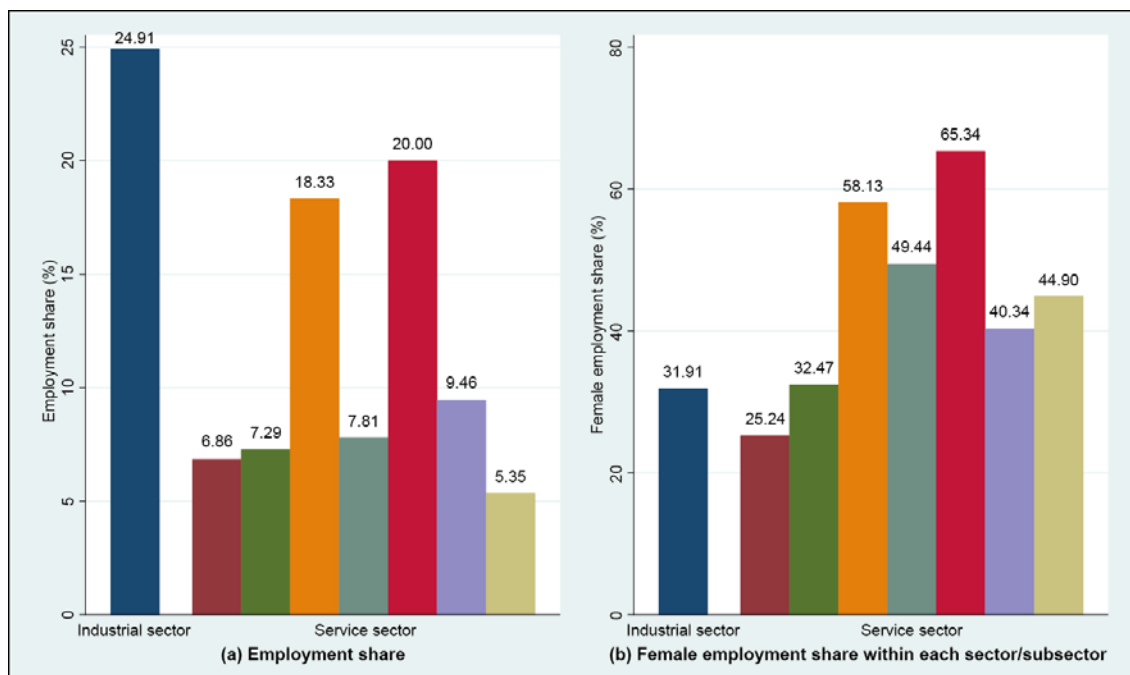
Figure 1: The Earnings Distribution by Gender and Distributional Gender Earnings Gap



Note: The sampling weights are used when calculating the summary statistics.
 Data source: CHFS 2017.

Figure 2 presents the employment shares of the industrial sector and service subsectors in the whole urban labor market, and the female employment shares within each sector/subsector. The industrial sector employs a quarter of urban workers approximately, and the service sector employs the other three quarters. About 30% of the workers in the industrial sector are men, while the female employment share in the service economy is about half, and the ratios in service subsectors 3 and 5, which also employ most workers among the seven service subsectors, are even as high as 58% and 65%, respectively.

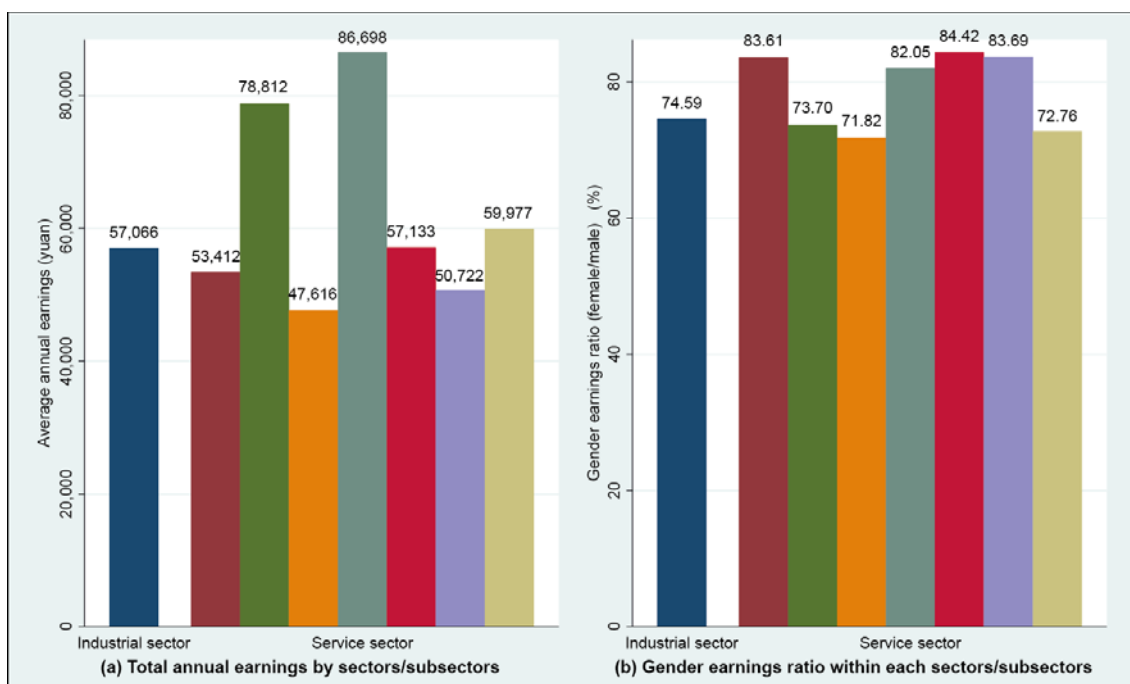
Figure 2: The Employment Shares by Economic Sector/Subsector and Female Employment Share Within Each Economic Sector/Subsector



Note: The sampling weights are used when calculating the summary statistics; see the text in Section 3.2 and the note of Table 1 for the classifications of service subsectors.

Data source: CHFS 2017.

Figure 3: The Average Annual Earnings by Economic Sector/Subsector and the Gender Earnings Ratios Within Each Economic Sector/Subsector



Note: The sampling weights are used when calculating the summary statistics; see note of Table 1 for the classification of service subsectors.

Data source: CHFS 2017.

Figure 3 presents the average annual earnings and gender earnings ratio in each sector/subsector. Figure 3(a) shows that on average the service sector has higher annual earnings than the industrial sector, but the heterogeneity is also serious. The average annual earnings in the finance and real estate industry are nearly twice those in the wholesale, retail, and accommodation and catering service industries. Figure 3(b) suggests that the gender earnings gap in the service sector is generally narrower than that in the industrial sector.

To sum up, the above descriptive analyses tell us that the service sector employs more women than the industrial sector, and the service sector has a narrow gender earnings gap. Therefore, these stylized facts seemingly suggest that when the service economy grows larger and larger, which should be the future trend of the PRC's economic development, the gender earnings gap should narrow. However, the extent to which the service economy matters needs more careful quantitative analyses.

3. METHODOLOGY

3.1 OB Decomposition

OB decomposition was separately developed by Oaxaca (1973) for the gender wage gap and Blinder (1973) for the black-white wage differential. Since then, the method has been widely applied in the labor market and discrimination studies, and the method itself has also been developed continuously to accommodate the issues arising in the empirical research.

In this study, let m and f denote male and female, respectively. The equation for male and female earnings is:

$$\ln Y_{gi} = X'_{gi} \beta_g + \varepsilon_{gi}, g = m, f, \tag{1}$$

where Y_{gi} is the annual earning of individual i , X_{gi} is a vector of explanatory variables including age, age squared, and dummies for educational attainment, Hukou and marital status, industry, occupation, and province where individual i resides. Under the usual assumptions $E[\varepsilon_{gi}|X_{gi}] = 0$, for $g = m, f$, the difference between the average earnings of males and females is:

$$\begin{aligned} \overline{\ln Y}_m - \overline{\ln Y}_f &= (\bar{X}'_m \hat{\beta}_m) - (\bar{X}'_f \hat{\beta}_f) \\ &= \underbrace{(\bar{X}_m - \bar{X}_f)' \hat{\beta}_f}_{\text{Composition effect}} + \underbrace{\bar{X}'_m (\hat{\beta}_m - \hat{\beta}_f)}_{\text{Structure effect}} \end{aligned} \tag{2}$$

$$\text{or} = \underbrace{(\bar{X}_m - \bar{X}_f)' \hat{\beta}_m}_{\text{Composition effect}} + \underbrace{\bar{X}'_f (\hat{\beta}_m - \hat{\beta}_f)}_{\text{Composition effect}}, \tag{3}$$

where $\hat{\beta}_m$ and $\hat{\beta}_f$ are the estimated coefficients for male and female samples, respectively; in fact, $\bar{X}'_m \hat{\beta}_f$ and $\bar{X}'_f \hat{\beta}_m$ are the two counterfactuals – the average male earnings if the male have the female coefficients of explanatory variables, and the average female earnings if the female have male coefficients to explanatory variables. Based on these counterfactuals, the overall gender earnings gap can be decomposed into “composition effect” and “structure effect.” The composition effect is attributed to the gender differences in the characteristics and thus often called the “endowment effect” or

the “explained”; the structure effect is often attributed to discrimination but generally it captures the effects of the gender differences in all of the unobserved variables (Jann 2008), and is thus also called the “unexplained.”

Comparing Eq. (2) and Eq. (3), we can see that the decomposition results depend on whose coefficients – males’ or females’ – are used as the reference when constructing counterfactuals. This is called the “index number problem” in the literature (Oaxaca 1973). Generally what coefficients should be used as the reference depends on the specific contexts and economic theory. Nevertheless, Reimers (1983), Cotton (1988), Neumark (1988), Oaxaca and Ransom (1994), Fortin (2008), and others proposed a series of solutions to avoid this problem. Here, I adopt Fortin’s (2008) solution: using the coefficients from a pooled sample regression with a gender dummy as the reference:

$$\overline{\text{Ln}Y_m} - \overline{\text{Ln}Y_f} = \underbrace{(\bar{X}_m - \bar{X}_f)' \hat{\beta}_p}_{\text{Composition effect}} + \underbrace{\bar{X}_m(\hat{\beta}_m - \hat{\beta}_p) + \bar{X}_f'(\hat{\beta}_p - \hat{\beta}_f)}_{\text{Structure effect}} \quad (4)$$

where $\hat{\beta}_p$ is the vector of the estimated coefficients of covariates in the pooled sample regression with a female dummy.

OB decomposition has another problem: For categorical variables, what group is omitted when generating dummy variables leads to different decomposition results (Jones 1983; Gardeazabal and Ugidos 2004; Yun 2005, 2008). Oaxaca and Ransom (1999) and Yun (2005) view this as an identification problem, while Gelbach (2002) believes this is not an identification problem but a conceptual problem. Fortin, Lemieux, and Firpo (2011) point out that there is no general solution to this problem, and their underlying conclusion is to omit the same group as previous studies to make the results comparable. It is fortunate that in this study, we are interested in the service economy and thus the industrial sector is the naturally omitted group; and for education and occupation, I find that the decomposition results do not change much when omitting different groups.

3.2 RIF Regression and RIF-OB Decomposition

The classical OB decomposition is very useful, but the mean is not the only statistic of interest. In regard to the gender earnings gap, people care about the mean gender gap, but the sticky floor and glass ceiling, as presented by Figure 1, are also important. However, the classical OB decomposition fails to provide any insight into them. In contrast, the RIF regression developed by Firpo et al. (2009) can solve this problem, and the RIF-OB decomposition (Firpo et al. 2007, 2018) can decompose the difference in any statistic between two groups. Specifically, for a random variable Y , the RIF is defined as:

$$\text{RIF}(y; v(F_Y), F_Y) = v(F_Y) + \text{IF}(y; v(F_Y), F_Y), \quad (5)$$

where $v(F_Y)$ is a distributional statistic of interest, and $\text{IF}(y; v(F_Y), F_Y)$ is the influence function of this specific distributional statistic. The influence function $\text{IF}(y; v(F_Y), F_Y)$ of the $v(F_Y)$ represents the influence of an individual observation on that statistic, and it is constructed with a zero mean. As a result, the RIF has a conventional feature:

$$v(F_Y) = E_Y[\text{RIF}(y, v(F_Y))].$$

If we introduce the explanatory variables and use the law of iterated expectation, we have:

$$v(F_Y) = E_X[E_Y[RIF(y, v(F_Y)|X)]] = E_X[m(x)],$$

where $m(x)$ is a function of X in any functional form. Furthermore, if we assume that $m(x)$ is a linear function, we have:

$$v(F_Y) = E_X[E_Y[RIF(y, v(F_Y)|X)]] = E[X]' \beta,$$

and the RIF regression:

$$RIF(y, v(F_Y)) = X\beta + \varepsilon, \quad (6)$$

where β measures the effects of the distributional change of explanatory variable on the distributional statistic $v(F_Y)$. One of the most important strengths of the above expression is that it “can be used directly for the estimation of standard errors of any statistic for which a RIF exists” (Rios-Avila 2020).

When the statistic $v(F_Y)$ is the quantile (of the dependent variable’s marginal distribution), RIF regression becomes unconditional quantile regression (Firpo et al. 2009). RIF-OB decomposition is a combination of RIF regression and the classical OB decomposition (Firpo et al. 2007, 2018). In my case, the estimate for the gender earnings gap can be expressed and decomposed as:

$$\hat{q}_\tau(F_Y^m) - \hat{q}_\tau(F_Y^f) = \underbrace{(\bar{X}_m - \bar{X}_f)' \hat{\beta}_p^\tau}_{\text{Composition effect}} + \underbrace{\bar{X}_m(\hat{\beta}_m^\tau - \hat{\beta}_p^\tau) + \bar{X}_f'(\hat{\beta}_p^\tau - \hat{\beta}_f^\tau)}_{\text{Structure effect}}, \quad (7)$$

where F_Y^m and F_Y^f are the cumulative distribution functions of the male and female logged earnings, respectively; $\hat{\beta}_m^\tau$, $\hat{\beta}_f^\tau$, and $\hat{\beta}_p^\tau$ are the unconditional quantile regression estimates for the coefficients of covariates at the τ th quantile for the male, female, and pooled samples, respectively. Eq. (7) is actually the quantile version of Eq. (3). In other words, the OLS and the classical OB decomposition are special cases of RIF regression and RIF-OB decomposition, respectively, since the statistic $v(F_Y)$ can be mean.

3.3 OB and RIF-OB Decomposition with Reweighting

OB decomposition based on a linear regression model is tractable and easy to understand. However, if the linear assumption does not hold, the decomposition results are likely to be biased (Barsky et al. 2002; Fortin, Lemieux, and Firpo 2011). One solution to this problem is reweighted regression. The idea is to make the explanatory variables in one group (for example, the males in this study) have similar distributions to those in the other group (the females in this study) by reweighting. Following DiNardo, Fortin, and Lemieux (1996), the reweighting function to make the males similar to the females in terms of explanatory variables is:

$$\Psi(X) = \frac{\Pr(g = f|X) \Pr(g = m)}{\Pr(g = m|X) \Pr(g = f)}.$$

In practice, the conditional probability is usually estimated using a logit or probit model, and the covariates for estimating the conditional probability can be the same as in the earnings equation, or sometimes a few interactions or polynomial terms of the covariates can be included.

Based on this reweighting function, the counterfactual coefficients and counterfactual mean can be estimated by

$$\hat{\beta}_c = \left(\sum \hat{\Psi}(X_{mi}) X_{mi} X_{mi}' \right)^{-1} \sum \hat{\Psi}(X_{mi}) y_{mi} X_{mi},$$

and

$$\bar{X}_c = \sum \hat{\Psi}(X_{mi}) X_{mi}.$$

Replace the counterfactual $\bar{X}_f \hat{\beta}_m$ with $\bar{X}_c \hat{\beta}_c$, and the decomposition in Eq. (3) changes to:

$$\overline{LnY}_m - \overline{LnY}_f = \underbrace{\bar{X}'_m \hat{\beta}_m - \bar{X}'_c \hat{\beta}_c}_{\text{Composition effect}} + \underbrace{\bar{X}'_c \hat{\beta}_c - \bar{X}'_f \hat{\beta}_f}_{\text{Structure effect}}, \quad (8)$$

where the composition and structure effects can be further decomposed as follows:

$$\text{Comp. effect: } \bar{X}'_m \hat{\beta}_m - \bar{X}'_c \hat{\beta}_c = \underbrace{(\bar{X}_m - \bar{X}_c) \hat{\beta}_m}_{\text{Pure comp.effect}} + \underbrace{\bar{X}_c (\hat{\beta}_m - \hat{\beta}_c)}_{\text{Specification error}};$$

$$\text{Struct. effect: } \bar{X}'_c \hat{\beta}_c - \bar{X}'_f \hat{\beta}_f = \underbrace{\bar{X}'_f (\hat{\beta}_c - \hat{\beta}_f)}_{\text{Pure struct.effect}} + \underbrace{(\bar{X}_c - \bar{X}_f) \hat{\beta}_c}_{\text{Reweighting error}}.$$

Here the pure composition and pure structure effects are the primary interest in decomposition analyses. Actually, the OB decomposition with reweighting can be viewed as the sum of two separate OB decompositions: the decompositions of the differences between the male and the counterfactual, and between the counterfactual and the female.

The RIF-OB decomposition has the same problem as the classical OB decomposition when the linear assumption does not hold. Similarly, I also conduct the RIF-OB decomposition based on reweighted regressions:

$$\hat{q}_\tau(F_Y^m) - \hat{q}_\tau(F_Y^f) = \underbrace{(\bar{X}_m - \bar{X}_c) \hat{\beta}_m^\tau}_{\text{Pure comp.effect}} + \underbrace{\bar{X}_c (\hat{\beta}_m^\tau - \hat{\beta}_c^\tau)}_{\text{Specif.error}} + \underbrace{\bar{X}'_f (\hat{\beta}_c^\tau - \hat{\beta}_f^\tau)}_{\text{Pure struct.effect}} + \underbrace{(\bar{X}_c - \bar{X}_f) \hat{\beta}_c^\tau}_{\text{Rew.error}},$$

where $\hat{\beta}_c^\tau$ is the reweighted RIF regression at the τ th quantile.⁴

In the following section, I conduct the classical OB and RIF-OB decompositions first, and then conduct the decompositions with reweighting. We will see whether our results are robust to reweighting or not.

⁴ Hereafter, τ th quantile and $(100 \cdot \tau)$ th percentile will be used interchangeably.

4. RESULTS

4.1 OLS and Unconditional Quantile Regression Results

Before looking into the decomposition results, let us look at and discuss a set of regression results, which are the bases of the decomposition analyses. Table 2 presents the OLS results (column 1) and unconditional quantile regression results (the remaining columns) for the log of earnings, and Panels A and B report the male and female results, respectively.

Table 2: The OLS and Unconditional Quantile Regression Results by Gender

Variables	OLS (1)	P5 (2)	P10 (3)	P25 (4)	P50 (5)	P75 (6)	P90 (7)	P95 (8)
Panel A: Male								
Service sector	-0.105*** (0.032)	-0.280*** (0.106)	-0.135** (0.061)	-0.052** (0.024)	-0.091*** (0.030)	-0.121*** (0.038)	-0.029 (0.063)	-0.032 (0.097)
Age	0.104*** (0.018)	0.324*** (0.091)	0.184*** (0.041)	0.073*** (0.014)	0.080*** (0.018)	0.074*** (0.019)	0.042* (0.025)	0.047 (0.036)
Age squared	-0.001*** (0.000)	-0.004*** (0.001)	-0.002*** (0.001)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001 (0.000)	-0.001 (0.000)
Educational attainment								
Senior high school	0.169*** (0.029)	0.478*** (0.179)	0.376*** (0.081)	0.167*** (0.035)	0.189*** (0.037)	0.145*** (0.038)	0.024 (0.048)	-0.003 (0.069)
Three-year college	0.320*** (0.033)	0.650*** (0.193)	0.438*** (0.098)	0.274*** (0.040)	0.342*** (0.041)	0.319*** (0.035)	0.198*** (0.065)	0.191* (0.101)
University and above	0.584*** (0.034)	0.677*** (0.195)	0.569*** (0.096)	0.406*** (0.043)	0.596*** (0.039)	0.683*** (0.048)	0.682*** (0.105)	0.743*** (0.165)
Hukou								
Nonagricultural	-0.018 (0.038)	0.022 (0.164)	-0.058 (0.079)	-0.055* (0.032)	-0.056 (0.034)	-0.031 (0.051)	0.100* (0.056)	0.193** (0.091)
Others	-0.063 (0.044)	-0.144 (0.189)	-0.108 (0.096)	-0.095** (0.040)	-0.110*** (0.042)	0.014 (0.058)	0.016 (0.070)	0.162 (0.137)
Single	-0.160*** (0.040)	-0.330* (0.190)	-0.279*** (0.079)	-0.118*** (0.035)	-0.108*** (0.039)	-0.124*** (0.043)	-0.252*** (0.079)	-0.360*** (0.115)
Occupations								
Profession	-0.040 (0.053)	0.121 (0.169)	0.026 (0.074)	-0.008 (0.032)	-0.019 (0.045)	-0.009 (0.064)	0.043 (0.090)	-0.293 (0.183)
Clerical staff	-0.180*** (0.057)	-0.129 (0.198)	-0.120 (0.080)	-0.101*** (0.039)	-0.138*** (0.046)	-0.140** (0.065)	-0.170* (0.092)	-0.548*** (0.194)
Service staffs	-0.221*** (0.060)	-0.225 (0.228)	-0.158 (0.109)	-0.143*** (0.048)	-0.166*** (0.059)	-0.150** (0.065)	-0.150* (0.085)	-0.433** (0.179)
Manufacturing workers	-0.274*** (0.076)	-0.527* (0.276)	-0.159 (0.123)	-0.151*** (0.050)	-0.229*** (0.068)	-0.270*** (0.074)	-0.190* (0.112)	-0.550*** (0.208)
Other occupations	-0.218*** (0.062)	-0.566* (0.337)	-0.380*** (0.143)	-0.185*** (0.054)	-0.085 (0.065)	-0.042 (0.086)	-0.080 (0.099)	-0.494** (0.221)
Provincial fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	9.012*** (0.321)	3.444** (1.688)	6.699*** (0.754)	9.136*** (0.262)	9.320*** (0.325)	9.977*** (0.350)	11.069*** (0.448)	11.919*** (0.659)
Observations	6,812	6,812	6,812	6,812	6,812	6,812	6,812	6,812
R-squared	0.232	0.056	0.094	0.140	0.187	0.169	0.120	0.088

continued on next page

Table 2 continued

Variables	OLS (1)	P5 (2)	P10 (3)	P25 (4)	P50 (5)	P75 (6)	P90 (7)	P95 (8)
Panel B: Female								
Service sector	-0.088** (0.035)	-0.289*** (0.099)	-0.106* (0.062)	-0.100** (0.046)	-0.090** (0.042)	-0.045 (0.067)	0.050 (0.063)	0.055 (0.067)
Age	0.074*** (0.016)	0.182** (0.070)	0.099*** (0.031)	0.088*** (0.021)	0.054*** (0.016)	0.045** (0.022)	0.051 (0.033)	0.044* (0.023)
Age squared	-0.001*** (0.000)	-0.002** (0.001)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.000 (0.000)	-0.001 (0.000)	-0.000 (0.000)
Educational attainment								
Senior high school	0.213*** (0.039)	0.380** (0.162)	0.321*** (0.090)	0.332*** (0.067)	0.189*** (0.036)	0.168*** (0.041)	0.074 (0.045)	0.078** (0.038)
Three-year college	0.497*** (0.043)	0.738*** (0.159)	0.597*** (0.090)	0.608*** (0.073)	0.489*** (0.048)	0.488*** (0.061)	0.276*** (0.062)	0.246*** (0.066)
University and above	0.727*** (0.042)	0.766*** (0.164)	0.670*** (0.094)	0.752*** (0.077)	0.683*** (0.042)	0.816*** (0.062)	0.751*** (0.097)	0.672*** (0.101)
Hukou								
Nonagricultural	-0.008 (0.035)	-0.138 (0.114)	0.054 (0.069)	0.003 (0.047)	-0.007 (0.035)	0.017 (0.044)	0.044 (0.067)	-0.033 (0.064)
Others	-0.058 (0.046)	-0.451*** (0.159)	-0.005 (0.072)	-0.053 (0.057)	-0.055 (0.036)	0.050 (0.059)	0.099 (0.086)	0.039 (0.091)
Single	-0.131*** (0.039)	-0.192 (0.130)	-0.269*** (0.071)	-0.158*** (0.059)	-0.114*** (0.039)	-0.064 (0.058)	-0.059 (0.097)	-0.163** (0.082)
Occupations								
Profession	-0.097 (0.067)	-0.033 (0.187)	-0.049 (0.081)	-0.035 (0.073)	-0.055 (0.081)	-0.010 (0.111)	-0.145 (0.173)	-0.329 (0.201)
Clerical staff	-0.218*** (0.063)	-0.078 (0.184)	-0.059 (0.081)	-0.133* (0.077)	-0.203*** (0.071)	-0.165 (0.114)	-0.302* (0.168)	-0.374* (0.210)
Service staff	-0.278*** (0.069)	-0.011 (0.210)	-0.049 (0.095)	-0.266*** (0.084)	-0.286*** (0.077)	-0.189* (0.107)	-0.395** (0.182)	-0.349 (0.228)
Manufacturing workers	-0.259*** (0.087)	-0.112 (0.271)	0.131 (0.126)	-0.078 (0.121)	-0.284*** (0.093)	-0.224* (0.134)	-0.389** (0.174)	-0.364* (0.205)
Other occupations	-0.204*** (0.077)	-0.102 (0.276)	-0.070 (0.120)	-0.145 (0.092)	-0.215** (0.095)	-0.101 (0.127)	-0.292 (0.184)	-0.334 (0.221)
Provincial fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	9.237*** (0.281)	5.942*** (1.411)	7.649*** (0.576)	8.396*** (0.396)	9.579*** (0.285)	10.084*** (0.397)	10.785*** (0.647)	11.331*** (0.530)
Observations	6,103	6,103	6,103	6,103	6,103	6,103	6,103	6,103
R-squared	0.249	0.038	0.082	0.168	0.227	0.198	0.116	0.067

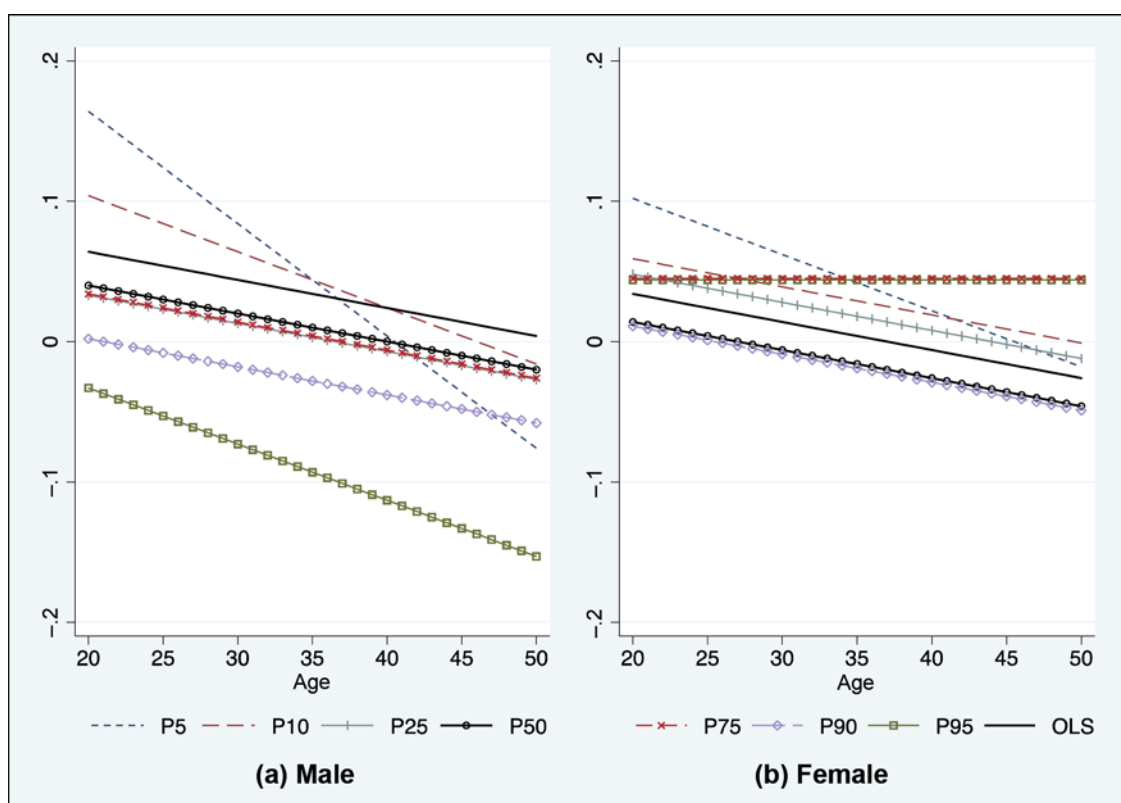
Note: Junior high school or below education, industrial sector, management job, and Beijing are the omitted categories in the regressions. Sampling weights are used in regressions. Robust standard errors in parentheses, which are clustered at the city level. Significance codes: *** p < 0.01, ** p < 0.05, * p < 0.10.

I begin with males (Panel A). First, the OLS estimate shows that a worker in the service sector earns 10.5% less than those in the industrial sector *ceteris paribus*. This forms a stark contrast with the descriptive analyses where workers in the service sector have higher annual earnings. However, they are not necessarily contradictory. This is because the high-pay industries usually need highly educated workers, and when the educational level is controlled for, the premium of the service economy should reduce and even disappear totally. The unconditional quantile regression results show that the impact of the service sector has a strong heterogeneity across earnings distribution, though it is always negative in all seven quantiles. Basically, the impact of the service sector has an “N” shape across the male earnings distribution, and it is statistically insignificant and close to zero at the top end. This illustrates that compared to the industrial sector, the

service economy is not a high-pay sector overall. This is likely due to the large dispersion of earnings within the service sector, which has been illustrated by Figure 3a and will be further discussed later.

I would like to point out that for the unconditional quantile regression results, the interpretation is a little tricky. For example, in column 2 the effect of the service sector means that if the service sector increases by 10% but all other things do not change, then the income distribution will change and the 5th percentile in the changed male income distribution will be 0.028 log points lower than that in the original income distribution. Those at the 5th percentile in the original income distribution are not necessarily the same as in the changed income distribution. In other words, the rank is not preserved when the earnings distribution changes.

Figure 4: The Marginal Effect of Age on Earnings by Gender



Note: The marginal effects of age at the 25th and 75th percentiles are nearly overlapped for males; the marginal effects of age at the 75th and 95th percentiles are nearly overlapped for females.

In terms of age, because age enters the earnings equation as a quadratic function, the marginal effect depends on age. Therefore, I plot the marginal effect of age on earnings in Figure 4, and Figure 4a is for males. On average, the earnings grow with age when people are young, but because the growth rate decreases with age and it becomes negative around the age of 52, from then on, the male earnings begin to decrease. The unconditional quantile regression results show heterogeneous effects. The earnings of the poorest people increase most quickly when they are young, but the growth rate also decreases most quickly. The marginal effects of age at the middle percentiles (25th, 50th, and 75th) of the male earnings distribution are the same or nearly the same. The richest workers (90th and 95th percentiles) have a negative growth rate even at the beginning of their career life, but since the marginal effect of age is statistically insignificant, it is

more likely to be the case that for the richest workers their earnings have nothing to do with age.

Compared to junior high school and below, people with senior high school, three-year college and university and above education earn 0.169, 0.320, and 0.584 log points more, respectively.⁵ These are consistent with our expectation. The distributional returns of senior high school and three-year college have an “inverse-N” shape across earnings distribution. In contrast, the distribution returns of university and above education has a “U” shape. In the 95th quantile, the premium of university and above education is the highest at 0.743 log points, which means the annual earnings of the highest educated are approximately 2.10 times those of the lowest educated.

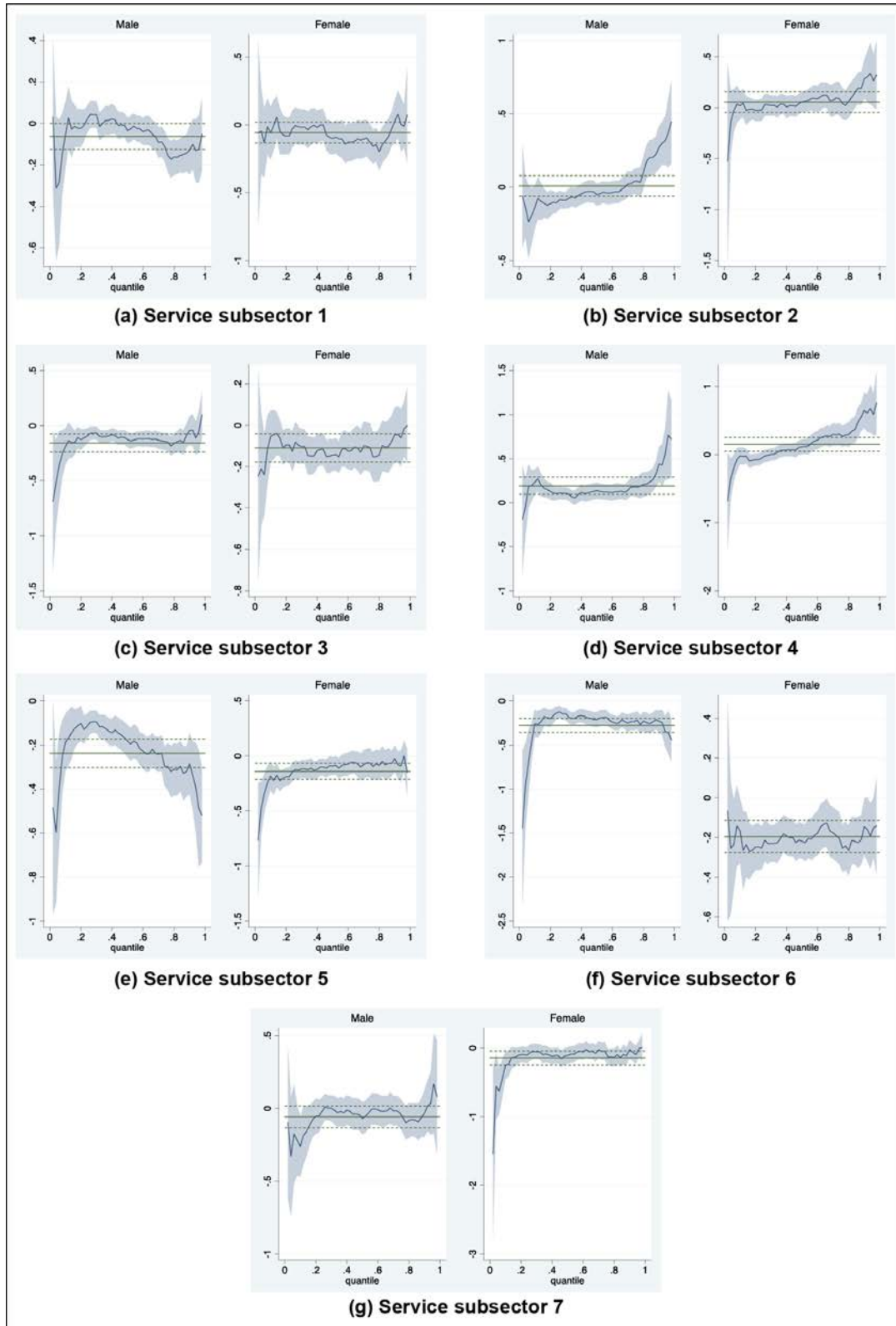
It is a little surprising that the nonagricultural Hukou has no premium relative to the agricultural Hukou on average. However, the unconditional quantile regression results suggest that the nonagricultural Hukou only has a statistically significant and positive premium in the 90th and 95th percentiles. To some extent, this demonstrates that labor market segregation still exists, but the Hukou system is not as important as before. The single people earn less than the others in all quantiles, especially at the top end of the earnings distribution. For occupations, the annual earnings of professional jobs have no significant difference compared to those of management jobs across the distribution. Nevertheless, all other occupations earn less than management jobs, especially at the 5th and 95th percentiles, though the premiums of management jobs relative to clerical and service staff are not statistically significant at the 5th and 10th percentiles.

For females (Panel B), the regression results are basically the same as those for males except for a few main differences. First, the negative premium of the service sector relative to the industrial sector monotonically decreases from the lower to the upper quantiles. Second, the average marginal effect of age on female annual earnings reduces to 0 by 37 years old, which is much quicker than for males (see Figure 4b). Third, in general, females enjoy a higher return rate to education with the exception of the highest educational attainment at the 95th percentile. On average, the return to education for females is at least 20% more than that for males. This implies that although females' education level is slightly higher than that of males, it is worthwhile investing in girls' education.

The descriptive analyses in the section on data show that there is a substantial heterogeneity in earnings and female employment shares among different service economic subsectors. To further investigate this heterogeneity, I substitute the dummy for the service sector with a set of dummies indicating the service economic subsectors. Meanwhile, the reference group is still the industrial sector. Figure 5 presents these regression results for the service economic subsectors. In the figure the horizontal lines represent the OLS results, and the broken lines and shaded areas indicate the unconditional quantile regression results. First, service subsectors 2 (logical, mathematical and engineering skill-intensive services) and 4 (finance and real estate) have strong heterogeneous effects across the earnings distribution. The expansion of these industries may widen the income inequality between the rich and poor, and may also widen the glass ceiling of women because rich men benefit more if these industries expand. Second, if subsector 5 (education, medical, social security and welfare) expands, people in all quantiles earn less, but the effect is not significant for females in the upper quantiles.

⁵ These correspond to approximately 18%, 38%, and 79%, respectively.

Figure 5: The Distributional Effects of Service Economy on Log of Annual Earnings



Note: The shaded areas are the 90% confidence intervals of the unconditional quantile regression estimates; the horizontal solid lines are the OLS estimates, and the horizontal dashed lines are the upper and lower bounds of the 90% confidence intervals of the OLS estimate. See note of Table 1 for the reclassification of service subsectors. Sampling weights are used in regressions.

Overall, these results suggest some heterogeneity in the effects of explanatory variables on earnings between males and females, and also some heterogeneity across different quantiles.

4.2 OB and RIF-OB Decomposition Results

Table 3 presents the basic OB and RIF-OB decomposition results, namely the decomposition results based on Eq. (4) and Eq. (7). Except for the industry composition, all other explanatory variables including age, age squared, and the other dummy variables, are grouped in “Age,” “Education,” “Occupation,” and “Province” to facilitate an interpretation of the results, and the contribution of the group is just the sum of the contributions of each variable contained in the group.

Table 3: OB Decomposition for the Mean Gender Earnings Gap and RIF-OB Decompositions at Different Quantiles

Variables	Mean (1)	P5 (2)	P10 (3)	P25 (4)	P50 (5)	P75 (6)	P90 (7)	P95 (8)
Male	10.802***	9.579***	10.045***	10.465***	10.780***	11.201***	11.688***	11.991***
Female	10.572***	9.477***	9.807***	10.136***	10.578***	11.047***	11.450***	11.733***
Raw gender gap	0.230***	0.102*	0.238***	0.330***	0.202***	0.154***	0.238***	0.257***
Composition effect	0.025**	0.072***	0.031**	0.021**	0.017*	0.010	0.010	0.021
Service sector	0.040***	0.075***	0.031**	0.021**	0.022**	0.026***	0.025**	0.042**
Subsector 1	-0.003	-0.007	-0.003	0.002	-0.002	-0.008***	-0.005	-0.005
Subsector 2	0.001	-0.004	-0.003	-0.003*	-0.001	0.002	0.011***	0.016***
Subsector 3	0.013***	0.028**	0.012*	0.009***	0.013***	0.014***	0.006	0.007
Subsector 4	-0.002*	0.001	-0.001	-0.001	-0.001	-0.003*	-0.006*	-0.007*
Subsector 5	0.028***	0.067***	0.031***	0.017***	0.018***	0.026***	0.023**	0.038***
Subsector 6	-0.005***	-0.011**	-0.005**	-0.004**	-0.004**	-0.005**	-0.004**	-0.006**
Subsector 7	-0.000	-0.001	-0.001	-0.000	-0.000	-0.000	-0.000	0.000
Age	0.002	0.005	0.002	0.001	0.002	0.002	0.001	0.001
Education	-0.009	-0.003	-0.003	-0.005	-0.008	-0.012*	-0.015**	-0.016**
Occupation	0.012***	0.005	0.012***	0.013***	0.010***	0.008**	0.016***	0.017**
Hukou	0.001	0.003	0.001	0.001	0.001	-0.001	0.001	0.000
Single	-0.005***	-0.009**	-0.009***	-0.004***	-0.004***	-0.003**	-0.006**	-0.009**
Province	-0.008**	-0.003	-0.002	-0.005*	-0.006*	-0.011***	-0.013**	-0.015**
Structure effect	0.204***	0.031	0.207***	0.309***	0.185***	0.144***	0.228***	0.236***
Service sector	-0.040	-0.033	-0.030	0.028	-0.012	-0.073	-0.085	-0.117
Subsector 1	-0.001	-0.009	-0.001	0.001	0.003	-0.002	-0.009	-0.002
Subsector 2	-0.003	0.001	-0.010	-0.006	-0.007	-0.004	-0.001	0.002
Subsector 3	-0.010	-0.039	-0.026	0.002	0.003	-0.012	-0.000	-0.018
Subsector 4	0.003	0.034*	0.020**	0.011*	0.002	-0.010	-0.017	0.003
Subsector 5	-0.019	0.007	-0.004	0.009	-0.016	-0.043**	-0.054**	-0.088***
Subsector 6	-0.008	-0.042*	-0.008	0.007	0.001	-0.002	-0.007	-0.023*
Subsector 7	0.005	0.015	-0.001	0.004	0.002	-0.001	0.002	0.010
Age	0.565	2.982	1.639*	-0.276	0.464	0.403	-0.149	0.072
Education	-0.093***	-0.058	-0.074	-0.237***	-0.046	-0.073*	-0.020	0.022
Occupation	0.018	-0.214	-0.089	0.029	0.069	0.037	0.193	-0.113
Hukou	-0.008	0.145	-0.084	-0.041	-0.040	-0.038	0.019	0.154**
Single	-0.007	-0.025	-0.003	0.007	-0.000	-0.014	-0.040**	-0.043*
Province	0.022	-0.181**	-0.146***	0.053**	0.027	0.053**	0.238***	-0.156***
Constant	-0.261	-2.587	-1.007	0.744	-0.276	-0.151	0.071	0.416
Observations	12,915	12,915	12,915	12,915	12,915	12,915	12,915	12,915

Note: See note of Table 1 for the classification of service subsectors. Junior high school or below education, industrial sector, management job, and Beijing are the omitted categories in the decompositions. Sampling weights are used in decompositions. The significance level is based on robust standard errors clustered at the city level. Significance codes: *** p < 0.01, ** p < 0.05, * p < 0.10.

Column (1) reports the OB decomposition results for the mean gender differential. It shows that the mean gender earnings gap is 0.23 log points, approximately equivalent to 26% of the female average earnings. Of this differential, 0.025 log points (11% of the mean gender earnings gap) can be explained by the differences in characteristics, and the difference in service sector contributes 160% to the total composition effect of the earnings gap. In other words, although the composition effect of the mean gender earnings gap is not large, the industry composition is the most important contributor among all explanatory variables. Within the service economy sector, subsectors 3 and 5 are the main contributors to widening the gender earnings gap. This is because these two service subsectors have lower average earnings than the industrial sector, but hire the most female workers. In other words, when female workers concentrate on low-pay industries, the gender gap should be widened, though the gender gaps within these industries are smaller than in other industries. Subsectors 1, 4, and 6 are helpful in narrowing the gender earnings gap, but the magnitudes are much less than in subsectors 3 and 5, and the contributions of subsectors 1 and 4 are insignificant in most quantiles. Occupation is another factor that widens the gender earnings gap, but the absolute contribution is no more than 30% of that of industry. Education and province make the same contributions to reducing the gender earnings gap, although education is not statistically significant. In the structure effect, age is the most important factor to widen the gender earnings gap. Remember that Figure 4 implies that the age-earning profile is quite different. In contrast, education is the main driver in narrowing the gender earnings gap. This is mainly because females have a higher rate of return to education. The service economy seems to be helpful in narrowing the gender earnings gap to some extent, but it is not statistically significant at all.

For the RIF-OB decomposition results at different quantiles, basically we find that the total gender earnings gap is lower in the lowest, middle, and upper middle quantiles. The share of the composition effect varies from quantile to quantile. The fraction of the composition effect in the raw gender gap is the highest and lowest in the 5th and 90th percentiles, respectively. In the 5th and 90th percentiles, about 71% and 4.2% of the raw gender gaps can be explained by the differences in characteristics. As for the service sector, it makes the largest contribution to the composition effect of the gender earnings gap in all quantiles. Within the service sector, subsector 3 (wholesale, retail, accommodation and catering service) mainly widens the gender gap in the lower quantiles, while subsector 4 (finance and real estate) mainly narrows the gender gap in the upper quantiles. Subsectors 5 (education, medical, public security and welfare) and 6 (public administration, social organization, and international organization) widen and narrow the gender earnings gap, respectively, across the whole distribution. More subsector 2 may strengthen the glass ceiling, while more subsector 4 weakens the glass ceiling. This is consistent with the regression results presented in Figures 4b and 4d. In the structure part, there is no consistent pattern across quantiles. With respect to the service economy sector, it is on the whole helpful in narrowing the gender gap and the size is basically the same as its composition effect, but not significant. Within the service, subsectors 4 and 5 increase and reduce the gender gap in the lower and upper quantiles, respectively.

Table 4 presents the decomposition results based on reweighted regressions. Basically, the results with reweighting are the same as those without. The only noticeable change is that the contributions of education and service economic subsectors 3 and 4 are more likely to be statistically significant in the composition effect.

Table 4: OB Decomposition for the Mean Gender Earnings Gap and RIF-OB Decompositions at Different Quantiles Based on Reweighting

Variables	Mean (1)	P5 (2)	P10 (3)	P25 (4)	P50 (5)	P75 (6)	P90 (7)	P95 (8)
Male	10.802***	9.579***	10.045***	10.465***	10.780***	11.201***	11.688***	11.991***
Counterfactual	10.782***	9.566***	10.001***	10.450***	10.814***	11.204***	11.679***	11.921***
Female	10.572***	9.477***	9.807***	10.136***	10.578***	11.047***	11.450***	11.733***
Raw gender gap	0.230***	0.102*	0.238***	0.330***	0.202***	0.154***	0.238***	0.257***
Total comp. effect	0.019**	0.023	0.044**	0.015*	-0.034***	-0.002	0.009	0.069**
Pure comp. effect	0.023**	0.052	0.021	0.007	0.014	0.017	0.014	0.043
Service sector	0.041***	0.069*	0.031*	0.016**	0.031***	0.046***	0.044***	0.079***
Subsector 1	-0.004	-0.010	-0.003	0.002	-0.002	-0.008***	-0.007*	-0.006
Subsector 2	0.000	-0.004	-0.006	-0.004**	-0.002	0.001	0.011**	0.016***
Subsector 3	0.014***	0.037*	0.017*	0.007*	0.011***	0.015***	0.004	0.010
Subsector 4	-0.002**	-0.002	-0.003**	-0.001**	-0.002**	-0.002**	-0.005***	-0.008**
Subsector 5	0.038***	0.062**	0.030**	0.015***	0.030***	0.045***	0.046***	0.075***
Subsector 6	-0.006***	-0.015**	-0.005**	-0.003***	-0.004***	-0.005***	-0.005***	-0.008**
Subsector 7	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	0.000
Age	-0.001	-0.004	-0.003	-0.001	-0.001	-0.001	-0.001	-0.001
Education	-0.013***	-0.003	-0.005	-0.007***	-0.013***	-0.018***	-0.023***	-0.027***
Occupation	0.012***	0.007	0.014*	0.010***	0.010***	0.009*	0.018***	0.026***
Hukou	0.000	0.001	0.001	0.001	0.001	-0.000	0.000	-0.001
Single	-0.006***	-0.012	-0.010***	-0.004***	-0.004**	-0.005**	-0.010***	-0.014***
Province	-0.010***	-0.006	-0.006*	-0.007***	-0.010***	-0.013***	-0.014***	-0.020***
Specification error	-0.004	-0.029	0.023	0.008	-0.048***	-0.020**	-0.005	0.026
Total struct. effect	0.210***	0.088	0.194***	0.314***	0.236***	0.157***	0.225***	0.192***
Pure struct. effect	0.200***	0.068	0.182***	0.307***	0.227***	0.149***	0.213***	0.179***
Service sector	-0.049	-0.117	-0.064	0.024	-0.020	-0.079	-0.120*	-0.172**
Subsector 1	-0.001	-0.012	-0.001	0.001	0.003	0.001	-0.008*	-0.004
Subsector 2	-0.005	-0.002	-0.012*	-0.007*	-0.007*	-0.004	-0.006	-0.006
Subsector 3	-0.010	-0.049	-0.018	0.007	0.007	-0.007	-0.016	-0.038
Subsector 4	0.002	0.028	0.019*	0.011	-0.000	-0.010	-0.022	-0.014
Subsector 5	-0.031*	-0.042	-0.035	0.002	-0.024	-0.058**	-0.062**	-0.098***
Subsector 6	-0.009	-0.052**	-0.015	0.007	-0.001	-0.003	-0.007	-0.020*
Subsector 7	0.004	0.013	-0.003	0.003	0.003	0.001	0.001	0.009
Age	0.429	3.980	1.951*	-0.107	-0.209	-0.128	-0.161	-0.280
Education	-0.101**	0.021	-0.016	-0.220***	-0.052	-0.080*	-0.066	-0.021
Occupation	0.051	-0.175	-0.055	0.035	0.076	0.022	0.242	0.001
Hukou	0.027	0.180	-0.062	-0.020	0.011	-0.012	0.063	0.188***
Single	-0.014	-0.020	-0.017	0.001	-0.013	-0.014	-0.037*	-0.027
Province	0.042**	-0.173*	-0.169***	0.064**	0.024	0.081***	0.277***	0.059
Constant	-0.184	-3.628	-1.386	0.530	0.410	0.360	0.016	0.429
Reweighting error	0.010	0.020	0.012	0.007	0.008	0.008	0.011	0.013
Observations	12,915	12,915	12,915	12,915	12,915	12,915	12,915	12,915

Note: See note of Table 1 for the classification of service subsectors. Junior high school or below education, industrial sector, management job, and Beijing are the omitted categories in the decompositions. Sampling weights are used in decompositions. The significance level is based on robust standard errors clustered at the city level. Significance codes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

To sum up, from the OB and RIF-OB decompositions, we find that the gender earnings gap in urban PRC in 2017 ranged from 10% to 33% approximately. Only about 10% of the gender gap can be explained by the differences in earnings equation's explanatory variables on average, and the industry composition can contribute to all or most of the composition effect. However, in the composition effect the expansion of the service sector is widening the gender earnings gap rather than narrowing it. This is mainly because, although the gender earnings gap is usually lower in the service sector,

females are more likely to work in the low-pay service subsectors and these subsectors also employ the majority of female workers. In terms of the structure effect, education has a strong narrowing effect on gender gap, though it is not significant over the whole distribution. As regards the service economy, the absolute contribution to the total gender gap in the structure effect is basically the same as that in the composition effect, but it is not statistically significant.

5. CONCLUSION AND DISCUSSION

Using the 2017 CHFS data, I investigate the gender earnings gap and the development of the service economy in urban PRC. The descriptive, regression, and decomposition analyses led to the following findings. 1) Compared to the industrial sector, the service economy hires more workers and has a higher female employment share, a lower gender earnings gap, and higher average annual earnings. 2) Females have a higher return to education rate than males generally, and holding other factors constant, the premium of the service sector relative to the industrial sector disappears and even becomes negative. 3) Only a small part of the total gender earnings gap is attributed to the composition effect, while the service economy contributes to all or most of the composition effect. 4) The service economy widens the gender earnings gap through the decomposition effect, but narrows the gender earnings gap through the structure effect, though the latter is not statistically significant.

Although more than half of the PRC's GDP has been created by a service economy, compared to the developed countries service economies still have much potential. Basically, there is no doubt that the development of the service economy is helpful in providing more jobs for females, and reducing the gender gaps in labor force participation. However, my results show that the development of the service economy does not necessarily narrow the gender earnings gap in urban PRC. The effect of the service economy on the gender earnings gap is complicated. This is because the service sector includes both high-pay industries like finance and real estate and low-pay industries like wholesale and retail, and accommodation and catering services, and females are more likely to work in the low-pay industries (see Figures 2 and 3). The barriers for women entering the high-pay industries can either be because of the public stereotype image of females or females' self-cognition about their abilities and roles in the society. The former is often related to discrimination, while the latter has been relatively less studied.

We have to admit that monetary income is only one concern of job choice, though it should be one of the most important factors. However, females may put more weight on family and choose to work in some low-pay industries. For example, females working in the education industry may earn less, but they have a stable job and enjoy their work-life balance and thus may obtain some premium in the marriage market. If this is true, females choosing to work in low-pay industries are more likely to have made a rational decision themselves rather than having been selected. This implies that to measure females' welfare we should comprehensively consider income, consumption, and even perhaps leisure. This should be the direction of future research about gender differentials.

REFERENCES

- Barsky, R., Bound, J., Charles, K., and Lupton, J., 2002. Accounting for the black–white wealth gap: A nonparametric approach. *Journal of the American Statistical Association* 97, 663–673.
- Blau, F. D., and Kahn, L. M. (2017). The gender earnings gap: extent, trends, and explanations. *Journal of Economic Literature* 55(3): 789–865.
- Blinder, A. S. (1973). Wage discrimination: Reduced form and structural estimates. *Journal of Human Resources* 8(4), 436–455.
- Borghans, L., ter Weel, B., and Weinberg, B. A. (2014). People skills and the labor-market outcomes of underrepresented groups. *ILR Review*, 67(2), 287–334.
- Ceci, S. J., Ginther, D. K., Kahn, S., and Williams, W. M. (2014). Women in academic science: A changing landscape. *Psychological Science in the Public Interest* 15(3), 75–141.
- Chi, W., and Li, B. (2014). Trends in China’s gender employment and pay gap: Estimating gender pay gaps with employment selection. *Journal of Comparative Economics* 42(3): 708–725.
- Cotton, J. (1988). On the Decomposition of Wage Differentials. *Review of Economics and Statistics* 70 (2): 236–43.
- DiNardo, J., Fortin, N. M., and Lemieux, T. (1996). Labor market institutions and the distribution of wages, 1973–1992: A semiparametric approach. *Econometrica* 64, 1001–1044.
- Fan, C. S., and Lui, H. (2003). Structural change and the narrowing gender gap in wages: Theory and evidence from Hong Kong. *Labour Economics* 10: 609–626.
- Firpo, S., Fortin, N. M., and Lemieux, T. (2009) Unconditional quantile regressions. *Econometrica* 77 (3), 953–973.
- Firpo, S., Fortin, N., and Lemieux, T. (2007). *Decomposing wage distributions using recentered influence function regressions*. Mimeo, University of British Columbia.
- Firpo, S., Fortin, N. and Lemieux, T. (2018). Decomposing wage distributions using recentered influence function regressions. *Econometrics* 6, 28.
- Fortin, N. M. (2008). The gender wage gap among young adults in the United States: The importance of money versus people. *Journal of Human Resources* 43(4), 884–918.
- Fortin, N., Lemieux, T., and Firpo, S. (2011). Decomposition methods in economics. In *Handbook of Labor Economics*, vol. 4A, edited by O. Ashenfelter and D. Card, 1–102. Amsterdam: Elsevier.
- Galor, O., and Weil, D. N. (1996). The gender gap, fertility, and growth. *The American Economic Review* 86 (3): 374–387.
- Gardeazabal, J., and A. Ugidos (2004). More on identification in detailed wage decompositions. *The Review of Economics and Statistics* 86(4), 1034–1036.
- Gelbach, J. B. (2002). Identified heterogeneity in detailed wage decompositions. Working paper, University of California at Berkeley.

- Goldin, Claudia. 2006. The quiet revolution that transformed women's employment, education, and family. *The American Economic Review* 96 (2): 1–21.
- Jann, B. (2008). The Blinder–Oaxaca decomposition for linear regression models. *The Stata Journal* 8(4), 453–479.
- Jones, F. L. (1983). On decomposing the wage gap: A critical comment on Blinder's method. *Journal of Human Resources* 18(1), 126–130.
- Kuhn, P., and Weinberger, C. (2005). Leadership skills and earnings. *Journal of Labor Economics* 23(3), 395–436.
- Lee, M. H. (2012). The one-child policy and gender equality in education in China: Evidence from household data. *Journal of Family and Economic Issues* 33(1), 41–52.
- Li, S., and Song, J. (2013). Changes in the gender-wage gap in urban China, 1995–2007. In *Rising Inequality in China: Challenges to a Harmonious Society*, edited by S. Li, H. Sato, and T. Sicular, 384–413. New York: Cambridge University Press.
- Li, S., Song, J., and Liu, X. (2014). The evolution of the gender earnings gap of the staff of China's cities and towns. *Management World* 03, 53–66.
- Li, Z., and Zhao, G. (2020). The quiet revolution in women's human capital and the gender earnings gap in the People's Republic of China. ADBI working paper No. 1112. <https://www.adb.org/sites/default/files/publication/579751/adbi-wp1112.pdf>.
- Liu, H. (2011). Economic reforms and gender inequality in urban China. *Economic Development and Cultural Change* 59(4), 839–876.
- NBS of China. (2020). Statistical Communiqué of the People's Republic of China on the 2019 National Economic and Social Development. February 28, 2020.
- Neumark, D. (1988). Employers' discriminatory behavior and the estimation of wage discrimination. *Journal of Human Resources* 23(3), 279–295.
- Ngai, L. R., and Petrongolo, B. (2017). Gender gaps and the rise of the service economy. *American Economic Journal: Macroeconomics* 9 (4): 1–44.
- Oaxaca, R. (1973). Male-female wage differentials in urban labor market. *International Economic Review* 14(3), 693–709.
- Oaxaca, R. L., and Ransom, M. R. (1994). On discrimination and the decomposition of wage differentials. *Journal of Econometrics* 61, 5–21.
- Pope, D. G., and Sydnor, J. R. (2010). Geographic variation in the gender differences in test scores. *Journal of Economic Perspectives* 24(2), 95–108.
- Reimers, C. W. (1983). Labor Market Discrimination against Hispanic and Black Men. *Review of Economics and Statistics* 65(4): 570–79.
- Rendall, M. P. (2017). Brain versus Brawn: The realization of women's comparative advantage. Institute for Empirical Research in Economics, University of Zurich, Working Paper No. 491, Revised version.
- Rios-Avila, F. (2020). Recentered Influence Functions (RIFs) in Stata: RIF Regression and RIF Decomposition. *The Stata Journal* 20 (1): 51–94.
- Song, J., Sicular, T., and Gustafsson, B. (2019). A review of China's urban gender wage gap from 1995 to 2013. *Japan Labor Issues* 3(17): 4–12.

- Tsui, M., and Rich, L. (2002). The only child and educational opportunity for girls in urban China. *Gender and Society* 16(1), 74–92.
- Weinberg, B. A. (2000). Computer use and the demand for female workers. *ILR Review* 53 (2): 290–308.
- Wu, L., Yan, K., and Zhang, Y. (2020). Higher education expansion and inequality in educational opportunities in China. *Higher Education*, 23 January.
- Yun, M.-S. (2005). A simple solution to the identification problem in detailed wage decompositions. *Economic Inquiry* 43(4), 766–772.
- . (2008). Identification problem and detailed Oaxaca decomposition: A general solution and inference. *Journal of Economic and Social Measurement* 33, 27–38.