The Gender Wage Inequality In Korea : A comparative analysis

Se Ri PARK*
Busan National University

Abstract

The main focus of this paper is analysing the adjusted gender pay gap in South Korea. To compare men and women similar in terms of observed characteristics – personal, educational, occupational and other job-related aspects – propensity score matching is applied together with the regression model and Blinder-Oaxaca decomposition. Using the 21^{st} wave of Korean Labor & Income Panel Study (KLIPS) surveyed in 2018, the results from the traditional regression-based model, Blinder-Oaxaca decomposition model, and a propensity score matching model were compared. The results revealed that first, the distribution of characteristics by gender is different and second, even after samples are restricted to those in common support, the unexplained component of the wage gap still seems significant. This implies not only the policy to increase the level endowments for women but also affirmative action to reduce the unexplained gap is required in South Korea to fight the gender wage inequality.

keywords: wage inequality, propensity score matching, gender pay gap
JEL Codes: J31, J71

* Ph.D Student, Graduate School of Economics, email:seri.park@pusan.ac.kr

I. Introduction

The gender equality has been one of major policy concern in Korea, but Korea is still far behind achieving its goal compared to its neighbourhood countries. Although Korean females are more educated than ever, their labour market outcomes remain poor, especially concerning labour force participation and wage. Female labour force participation rate has remained approximately 50% since 2010, and gender pay inequalities continue to be significant in South Korea.

Table 1. Economically Active Population and Labor Force Participation Rate for Women [unit : in 000, %].

	2010	2011	2012	2013	2014	2015	2016	2017	2018
Aged over 15	20,846	21,119	21,356	21,576	21,806	22,018	22,205	22,357	22,484
Economically active population	10,335	10,520	10,704	10,862	11,229	11,426	11,583	11,773	11,893
Labor force participation rate	49.6	49.8	50.1	50.3	51.5	51.9	52.2	52.7	52.9

KOSIS(2020). Economically Active Population Survey. (Accessed on 28 June 2020)

OECD (2020) defines the gender wage gap as “the difference between median earnings of men and women relative to median earnings of men”. According to the OECD (2020) gender wage gap indicator shown in figure 2, despite its decreasing trend, the gender wage gap in South Korea was recorded 34.1% in 2018 which was the highest amongst OECD countries. In the same year, the average gender wage gap of OECD member countries was 13.2%.

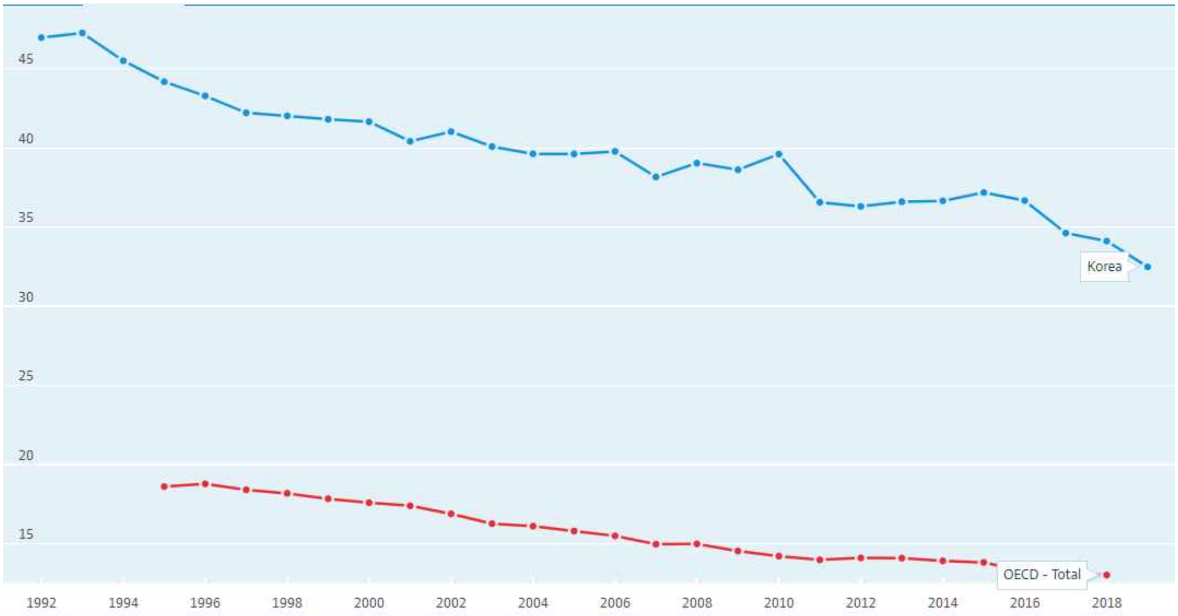

Figure 1. OECD Data on Gender Wage Gap

Statistic provided by Korean Statistical Information Service(KOSIS), shown in table 2 also shows the similar result. The average monthly wage for men was 3,135,000 (KRW), and that of female was 2,087,000 (KRW) in 2018. The female to male earnings ratio was 66.6 per cent, which indicates that female workers earn about 66.6 per cent of what male workers earn on average.

Table 2. Income differences by Gender [unit : in 000, %].

		2010	2011	2012	2013	2014	2015	2016	2017	2018
Female	Monthly Wage	1,477	1,548	1,654	1,705	1,742	1,781	1,869	1,946	2,087
	As % of male	62.6	63.3	64.4	64	63.1	62.8	64	64.7	66.6
Male	Monthly wage	2,361	2,444	2,569	2,664	2,761	2,837	2,918	3,010	3,135

*Female to male earnings ratio = (Monthly wages of females / monthly wages of males)*100*

KOSIS(2020) Labor Conditions by Employment Type (Accessed on 28 June 2020)

The purpose of this study is to estimate and compare the effect of being female on average monthly wages in Korea with parametric and semi-parametric methods. The most recent data available, the 21st wave of Korean Labor & Income Panel Study (KLIPS), is utilised to analyse the wage differentials by gender. The empirical setting of this work is similar to the work of Goraus et al. (2017) and Djurdjevic & Radyakin (2005). The simplest regression model is used to analyse the factors affecting the average wages of men and female. Then Blinder-Oaxaca decomposition analysis to decompose is applied to break the gender wage gap into two components; one that is explained by differences in the level of individual characteristics, and another that cannot be explained. This unexplained component is, sometimes, attributed to discrimination. But this approach suffers from the heterogeneity of data and sample selection bias. Hence the semi-parametric method is applied to compare those who are more homogeneous in their characteristics. Propensity score matching originated from statistical literature, is known to be a possible solution to these problems. To compare like with like, a matched sample of men and women who have similar socio-economic characteristics is created using the propensity score. Then, the results from the decomposition of wage gap with matching are compared to that of the traditional parametric decomposition to answer the question: how much of gender wage gap cannot be explained by differences in characteristics?

To best of my knowledge the importance of recognizing the problem of gender differences in the supports has not yet been addressed intentively in the studies on the gender wage gap in Korea. This paper differs from propensity score matching analysis adopted by Ahn & Oh (2017) who used a sample of individuals who were students or unemployed from 2007 Youth Panel to see whether their wages showed meaningful differences after seven years. With B-O decomposition and propensity score matching, they found the wage difference of 27.5% for men and women who are alike in terms of human endowment and job aspects. Still, 22.7% of dissimilarities were left unexplained even though they were alike at the early stage of career.

The rest of this paper is organised as follows: Chapter 2 represents the theoretical background and empirical literature on the gender pay gap in Korea. In chapter 3, a description of the dataset is provided followed by chapter 4 with the introduction of empirical methodology and results: Ordinary Least Squares (OLS) estimates of Mincerian wage function, B-O decomposition and propensity score matching. Conclusion in chapter 5 complies the result and discussion of the paper.

II. Review of Literature

1. Theoretical literature

Economic analysis of the causes of the gender wage gap have centred on gender differences in individual endowments, i.e., the human capital model or labour market treatment.

The human capital theory implies that differences in market-rewarded qualifications by gender account for the income inequalities. It suggests that an individual's incentive to invest in these qualifications, which will be translated to higher income in the future, are positively related to anticipated labour force participation period. Regarding costs and benefits to obtain marketable training, the longer one's lifetime work years, the higher the gain from the investment, ceteris paribus (Polachek, 2004). Women who expect, on average, shorter work lives and prospective discontinuity due to raising children and the role in the family have lower incentives to acquire human capital (Mincer and Polachek, 2004; Becker, 1985). This results in smaller investment in market-rewarded training which leads to lower wages compared to those of men.

Various form of labour market discrimination is further assumed to affect the earnings of women. One is taste-based discrimination model developed by Becker (1971). *In The Economics of Discrimination*, he introduced "a taste for discrimination", which means there is disamenity for hiring a particular type of worker. In his model, three types of discrimination are described: employer discrimination, co-worker discrimination and customers discrimination. Grybaite (2006) made the great examples of these discrimination : employers looking for engineers may not be willing to employ women but only men; Men who avoid working with the female boss may willing to hire a female secretary; Consumers unwilling to buy automobiles from female automobile dealers would like to shop cosmetics from female shop assistances. Because of the taste for discrimination, women as minority workers must compensate employers either by being more productive at a given wage or by accepting a lower salary for equal productivity.

Bergmann's (1971) crowding model is another theory based on discrimination. Women's limited access to some professions because of employer discrimination causes occupational segregation by gender. As they are crowded into specific jobs, this would result in the excess supply of female labour which lowers the average level of women's earnings. The occupational segregation by gender can explain why wages in female-dominated occupations are likely to be lower than those in male-dominated professions.

The statistical theory of discrimination proposed by Phelps (1972) and Arrow (1973) is the other standard formulation of discrimination (Guryan & Charles, 2013). The statistical discrimination model proposes that firms have limited information about the productivity of applicants. In the absence of perfect information, they would try to predict the expected productivity of applicants using easily observable characteristics such as ethnicity or gender to figure out the expected productivity of applicants. If they believe that women are generally less productive or less stable workers, an individual woman could be discriminated since they judge her to be less productive or less stable.

2. Empirical literature

A number of studies have focused on the gender wage gap in the Korean labour market. They attempted to find the existence and causes of the gender wage gap using Blinder-Oaxaca decomposition focusing on various factors. Introduced by Blinder (1973) and Oaxaca (1973), B-O decomposition has been one of the most extensively utilised methods in the wage gap literature. It compares the mean income differences in two groups, breaking the income differences into two parts: "explained part" or "endowments effect" and "unexplained part" or "coefficients effect". The majority of existing literature on the gender wage gap concluded that the gender wage gap in Korea attributed to the unexplained part or discrimination is acknowledging. Still, the magnitude of it varied due to differences in the data adapted and the specification of the wage equations.

Heo (2003) analysed the gender wage differentials in industry and job level by applying the B-O decomposition method used by Fields and Wolff (1995) on the 1999 Korean Labor & Income Panel Study (KLIPS). He found that in industry level and occupation level wage discrimination plays a substantial role compared to employment discrimination in explaining gender wage gaps which are affected by differences in productivity. It was also revealed that gender wage gaps are affected significantly by differences in productivity showing only 3.6-6.8% accounted for the unexplained part.

Using the 2001 Survey report on labour conditions by employment type, Kim (2013) examined the earnings gap based on both gender and employment type for a sample of 793,036 individuals. Oaxaca and Ransom's (1994) two-fold decomposition method was used in his analysis, concluding 44.9% of the gender wage gap and 18.9% of the total pay gap between regular and irregular worker were attributed to the unexplained part.

Shin (2011) employed the Blinder-Oaxaca decomposition with bootstrap method, using the Aditional Survey of the Economically Active Population 2007, to examine the structure of gender earning gap. He found that although rewards for job tenure and working hours are higher for women, women aged in the 40s and 50s experienced the biggest gap in their earnings compared to those of men. Further, approximately 44.2% of the gender wage gap was explained by the difference in endowments but 65% of it was left to unexplained.

In her article, *Effect of Marriage and Infant Child on Gender Wage Gap*, Jang (2020), employed the Blinder-Oaxaca decomposition combined with recentred influence function on the 2016-2018 pooled data of Local Area Labour Force Survey to analyse the effect of marriage and child-care on gender pay gap. For the observed wage differences between male and female, she found the unexplained wage gap of 56.3% for the single, of 61.98% for the married without kids and of 63.8% for the married with kids.

Yoo (2017), however, argued the discrimination against women prevailed even in the earlier stage of their lives. By utilising the Blinder-Oaxaca decomposition method referring to the sample of graduates in 2013 from the Graduates Occupational Mobility Survey

(GOMS), She analysed the gender earning differences in the young before career disruption of women occurs and found that differences in the rewards explained 32.1% of gender earning differences. One of the most interesting results from her work was that the average monthly incomes of women with bachelor’s degree were even less than those of men who graduated junior college, which proved gender effect had a bigger magnitude on the wage of women than the level of education.

III. Data and Descriptive Overview

1. Data

The data used in this study is the Korean Labor & Income Panel Study (KLIPS), which is a longitudinal survey of households and individuals living in urban areas. The Korean Labor Institute has published the detailed data for households and individuals covering employment, income, health and other various regions since 1998. Its design and management follow the Panel Study of Income Dynamics (PSID) directed by the University of Michigan and is comparable to other longitudinal surveys such as the German Socio-Economic Panel (SOEP) of Germany and the Survey of Labour and Income Dynamics (SLID) of Canada. From the 21st wave data surveyed in 2018, 4,298 individuals who are currently employed and reported their average monthly income are selected after those with missing variables for any explanatory variables are removed.

The sample of 2,512 males and 1,786 females wage earners aged between 19 and 60 are selected. In the KLIPS, the average monthly wage of individuals is reported in 10,000 Korean Won (KRW). The average monthly salary for the sample is 2.8 million KRW, which ranges from 150,000 KRW to 30,000,000 KRW. The level of education is measured with the number of years of schooling. The KLIPS, however, only provides the level of education as a categorical variable with a separate variable which represents the level of attainment. Hence, the education variable in this study is processed as follows: as primary school is compulsory education in Korea, all individuals are assumed to receive at least six years of education; those who have not finished the reported the level of education are assumed to achieve at least the level below; those who finished university or higher level of education are expected to receive 21 years of schooling.

The dummy variables are created for a residential area, industry, the size of the firm, occupation, marital status, the regularity of job and gender as represented in table 2. Industry variables are categorized using the Korean Standard Industrial Classification

(KSIC)-10. While primary industries are combined as one category (agriculture, forestry, fishing, mining and quarrying), three categories of (1) public administration and defence; compulsory social security, (2) Activities of households as employers; undifferentiated goods-and services-producing activities of households for own use and (3) activities of extraterritorial organizations and bodies have been removed. Further, these are sorted into primary, secondary, tertiary and quaternary industries to obtain a sufficient number of samples in each category. For occupation dummy, the International Standard Classification of Occupations (ISCO)-8 based on skill levels is used for categorising individuals into either high-skill job or low-skill job. High-skill jobs include managers and professionals (ISCO 1-3) whereas low-skill duties include clerical support workers, service and sales workers and elementary occupations (ISCO 4-8).

Table 3. Definition and Description of Variables

Data	2018 wage earners
Age	19-60
Region	All the regions
Industy	Exclude
	(1) Public administration and defence; compulsory social security
	(2) Activities of households as employers; undifferentiated goods-and services-producing activities of households for own use
	(3) Activities of extraterritorial organizations and bodies
Occupation	Classified based on KSCO matched with ISCO-8 based on skill levels
	Exclude the unemployed and soldiers
Variables	
lnwage	Log monthly wage
gen	Gender (female = 1)
age	Age
ft	Regularity of employment (regular workers =1)
edu	Years of schooling
	Those who have not finished the reported level of education are assumed achieve the level below
exp	Experience
	Survey year - employed year
exp2	Experience squared
mar	Marital status (Married = 1)

area	Regions (Grater Seoul = 1)
union	Union (Union exists = 1)
size	Size of firms in terms of the number of employees s1 : 1-30 employees s2 : 30-300 employees s3 : 300 employees or more
job	Job characteristics based on skill levels High skill (ISCO 1-3) = 1 Low skill (ISCO 4-8) = 0
ind	Industry characteristics ind1 : primary ind2: secondary ind3: tertiary ind4 : quaternary

2. Descriptive Overview

The main characteristics of the sample are represented in table 4. The average age of males and females in the sample is 41.524, whereas the average years of experience and education are 6.729 years and 13.763 years separately. The individuals on the sample data earn, on average, 2,857,200 KRW in a month.

Table 4. Summary Statistics of the data set

	Mean	STD	Min	Max
Age	41.524	10.187	19	60
Education	13.763	2.422	6	21
Monthly wage (0,000)	280.572	173.304	15	3000
Log monthly wage	5.489	0.551	2.708	8.006
Experience	6.729	7.119	0	41
Experience Squared	95.951	186.806	0	1681
Observations		4,298		

Table 5 describes the socio-economic aspects by gender. The mean monthly wage of men is 3,369,940 KRW while that of women is only 2,012,150 KRW.[2] The difference of

1,357,790 KRW is significant given the minimum hourly wage in 2018 was 7,530 KRW, translated to 1,573,770 KRW monthly.[3)]

The years of education and experience show no significant difference in terms of gender. The difference in experience years of approximately two could be explained by mandatory army services only served by men in Korea. But, the difference in marital status by gender shows statistical significance with a p-value less than 0.001 and χ^2 of 16.045. 36.75% of women are single, whereas 30.89% of men are not married.

In terms of occupation, over 70% of men and women are regular workers, but most of them have low-skill jobs. Only 26.91% of men and 30.40% of women are employed as high-skilled workers. The existence of union in establishments significantly differs by gender, reporting χ^2 of 96.458 ($p<0.001$). Approximately 90% of female employees work at firms without the union, while the proportion drops below 80% for male workers. It further exhibits the notable difference in the distribution of males and females concerning the size of the company. Men are more likely to work in medium-large size firms in terms of the number of employees.

In industrial aspect, the number of men working in secondary industry, mostly manufacturing, outnumbers that of women, by contrast, more women are working in tertiary sector, for example, education. Industrial segregation by gender is related to traditional social norms in Korea and ‘stigmatization’: blue-collar jobs are not only considered unsuitable for women in Korea but also society ‘stigmatize’ the spouses of female blue-collar workers (Lee et al., 2008)

---

2) The reason for using the average monthly wage in this paper rather than the average hourly wage is that the former represents the quality of employment better than the latter.

3) Miminum Wage Commission Republic of Korea, http://minimumwage.go.kr/eng/sub04.html

Table 5. Summary Statistics by gender

	Male		Female	
	Mean	STD	Mean	STD
Age	41.887	9.565	41.012	10.986
Education	14.044	2.403	13.368	2.395
Monthly wage (0,000)	336.994	183.053	201.215	119.927
Log monthly wage	5.706	0.482	5.184	0.493
Experience	7.864	7.705	5.132	5.842
Experience Squared	121.195	212.444	60.447	135.498
Observations	4,298			

	Male			Female			
	Freq.	Prop. (%)	Ave. Wage	Freq.	Prop. (%)	Ave. Wage	χ^2
Greater Seoul	1,275	50.76	350.721	953	53.36	207.802	2.833
Other areas	1,237	49.24	322.846	833	46.64	193.680	(0.092)
Married	1,736	69.11	377.042	1,130	63.27	201.810	16.045***
Single	776	30.89	247.403	656	36.73	200.191	(0.000)
High-skill job	676	26.91	398.678	543	30.40	221.912	6.266*
Low-skill job	1,836	73.09	314.282	1,243	69.60	192.174	(0.012)
1-30 workers	1,074	42.75	268.790	1,044	58.45	174.455	102.938*** (0.000)
30-300 workers	763	30.61	326.406	434	24.3	212.719	20.640*** (0.000)
300 and above workers	669	26.63	458.659	308	17.25	275.711	52.366*** (0.000)
Union	508	20.22	446.929	164	9.18	296.201	96.458***
No union	2,004	79.78	309.126	1,622	90.82	191.611	(0.000)
Regular job	2,157	85.87	356.097	1,330	74.47	223.368	88.611***
Irregular job	335	14.13	220.924	456	25.53	136.601	(0.000)
Primary	23	0.92	308.478	7	0.39	101.429	4.130* (0.042)
Secondary	1,227	48.85	348.588	330	18.48	207.552	416.693*** (0.000)
Tertiary	935	37.22	304.068	1,167	65.34	193.786	330.323*** (0.000)
Quaternary	327	13.02	389.642	282	15.79	227.021	6.595* (0.010)
Total	2,512	58.45	336.994	1,786	41.55	201.215	

p-value in parentheses

** p < 0.05, ** p < 0.01, *** p < 0.001*

Ⅳ. Methodology and Results

1. Mincer wage equation

The original Mincer (1974) earnings equation based on the human capital theory models the natural log of wage as a linear function of education and a quadratic function of potential experience. It suggests the human capital factors - years of scholling and potential labour market experience are the primary wage determinants.

$$\ln W_i = X_i^{'}\beta + \epsilon_i \tag{1}$$

Whereas $\ln W_i$ represents the natural log of average monthly income variable, $X_i^{'}$ is a matrix of control characteristics that possibly affect one's earning. To analyse the impact of social-economic characteristics on wages, the specification of the wage function used in this paper is:

$$\begin{aligned}\ln W_i &= a + \beta_1 edu_i + \beta_2 age_i + \beta_3 area_i + \beta_4 job_i + \beta_5 ft_i + \beta_6 mar_i \\ &+ \beta_7 gen_i + \beta_8 union_i + \beta_9 exp_i + \beta_{10} exp_i^2 + \sum_k \tau_k ind_{ki} + \sum_m \delta_m size_{mi} + \epsilon_i\end{aligned} \tag{2}$$

where α is an intercept for each individual, the details of other variables are shown in table 1. The wage gap between male and female workers controlling for differences in the personal-occupational specifications is measured by β_7.

Table 5 summarises the results of regression estimation by sex and pooled sample. It shows women's salaries tend to be 33.7% lower than the wages of men. While age and marital status have a positive effect on the wages of men but a negative effect on those of women, it is statistical significance only on males. The number of years of work experience and education have a positive impact on both gender, showing statistical significance at 1%. Besides, the skill level of job and firm size have a significant effect on the average monthly wages for both women and men. Workers who work at a firm with 300 and above employees would receive higher wages compared to those who work in smaller workplaces. The positive firm size effect on the average wage is observed in other industrialised countries such as Japan and Britain (Daly et al., 2006). In respect to primary industry, working in secondary or tertiary industries would increase the income by 7.9%-17.6% for male workers and by 1.3%-15.9% for female workers.

Table 6. Estimated coefficients from Mincer log earnings regression

	All	Males	Females
	lnwage	lnwage	lnwage
Gen	-0.337***		
	(0.013)		
Age	0.002*	0.004***	-0.001
	(0.001)	(0.001)	(0.001)
Edu	0.041***	0.040***	0.034***
	(0.003)	(0.003)	(0.005)
Exp	0.021***	0.017***	0.020***
	(0.002)	(0.003)	(0.004)
Exp2	-0.000	-0.000	0.000
	(0.000)	(0.000)	(0.000)
Mar	0.094***	0.195***	-0.024
	(0.014)	(0.018)	(0.021)
Area	0.039***	0.055***	0.018
	(0.011)	(0.014)	(0.018)
Ft	0.350***	0.299***	0.367***
	(0.016)	(0.023)	(0.022)
Job	0.091***	0.093***	0.081***
	(0.014)	(0.018)	(0.023)
Union	0.068***	0.055*	0.108**
	(0.019)	(0.022)	(0.037)
1-30 employees	-0.245***	-0.271***	-0.198***
	(0.018)	(0.022)	(0.030)
30-300 employees	-0.149***	-0.185***	-0.078*
	(0.018)	(0.021)	(0.032)
300 and above employees	0.000	0.000	0.000
	(.)	(.)	(.)

Primary	0.000	0.000	0.000
	(.)	(.)	(.)
Secondary	0.049***	0.079***	0.013
	(0.015)	(0.017)	(0.027)
Tertiary	0.188***	0.176***	0.159*
	(0.022)	(0.023)	(0.070)
Quaternary	0.004	0.045	-0.031
	(0.018)	(0.024)	(0.028)
Constant	4.606***	4.499***	4.476***
	(0.057)	(0.070)	(0.089)
N	4298	2512	1786

Standard errors in parentheses

** p < 0.05, ** p < 0.01, *** p < 0.001*

Reference group for dummy variables are S3 (300+ employees) and Ind1 (primary industry)

2. Blinder-Oaxaca decomposition

The gender wage gap is measured by the mean wage differences between males and females using the counterfactual decomposition proposed by Blinder (1973) and Oaxaca (1973). The Blinder-Oaxaca decomposition[4] breaks the wage gap into “explained” and “unexplained” components. Where the “explained” part represents the wage differences due to different productivity characteristics held by male and female, the “unexplained” part describes the discrimination against women or the effect of unobservable wage determinants (Jann, 2008).

The wage function of men and women as described as follows:

$$\ln W_M = W_M^{'}\beta_M + \epsilon_M$$
$$\ln W_F = W_F^{'}\beta_F + \epsilon_F$$

4) The theoretical explanation and equation for Blinder-Oaxaca decomposition in this paper follows Jann (2008).

where the subscripts of M and F are used to represent male and female, respectively. The difference in mean outcome becomes

$$D = E(\ln W_M) - E(\ln W_F) = E\left(X_M^{'}\right)\beta_M - E\left(X_F^{'}\right)\beta_F$$

assuming $E(\beta_M) = \beta_M$, $E(\beta_F) = \beta_F$, $E(\epsilon_F) = 0$ and $E(\epsilon_M) = 0.$

(3)

By rearranging (3), the difference in the mean outcome can be divided into three terms:

$$D = \{E(X_M) - E(X_F)\}^{'}\beta_M + E\left(X_M^{'}\right)(\beta_M - \beta_F) + \{E(X_M) - E(X_F)\}^{'}(\beta_M - \beta_F)$$

(4)

Given the ordinary least-squares estimates $\widehat{\beta_M}$ and $\widehat{\beta_F}$ and mean wages of males and females $\overline{X_M}$ and $\overline{X_F}$, equation (4) can be estimated:

$$\widehat{D} = \overline{Y_M} - \overline{Y_F} = (\overline{X_M} - \overline{X_F})^{'}\widehat{\beta_M} + \overline{X_M}^{'}(\widehat{\beta_M} - \widehat{\beta_F}) - (\overline{X_M} - \overline{X_F})^{'}(\widehat{\beta_M} - \widehat{\beta_F})$$

(5)

The first term shows the 'endowments effect' which is the wage differential driven by differences in the productivity-related characteristics captured by independent variables. It effectively represents the expected change in the wage of a male if they had the same attributes as female. The second term refers to the 'coefficients effect' meaning the expected change in wage of men when coefficients of women are applied. The last term is the interaction term; that is, the simultaneous effect of endowment and coefficients.

With an assumption of a nondiscriminatory coefficient vector β^*, it is possible to divide the outcome difference D, into two terms. If there is a nondiscrimination coefficient vector β^*, it is possible to show the equation (4) as follows

$$D = \{E(X_M) - E(X_F)\}' \beta^* + \{E(X_M)'(\beta_M - \beta^*) + E(X_F)'(\beta^* - \beta_F)\} \tag{6}$$

In this "twofold" decomposition model, the first term is the quantity effect that can be explained by the difference in wage determinants held by men and women. The second part, commonly named as "discrimination effect", is the difference in the wage that cannot be explained by the variation in the predictors but can be attributed to differences in the coefficients.

The various suggestions on the estimate of β^* have been made in the literature. Oaxaca (1973) assumed that discrimination is directed toward one group, not the other one, $\beta^* = \beta_M$ or $\beta^* = \beta_F$, therefore,

$$\hat{D} = (\overline{X_M} - \overline{X}_F)' \widehat{\beta}_M + \overline{X_F}' (\widehat{\beta}_M - \widehat{\beta}_F)$$

or

$$\hat{D} = (\overline{X_M} - \overline{X}_F)' \widehat{\beta}_F + \overline{X_M}' (\widehat{\beta}_M - \widehat{\beta}_F)$$

Cotton (1988) argued that the average coefficient from two groups should be used for the estimation of β^* because Oaxaca's (1973) assumption could lead to the overvaluation of one group and the undervaluation of the other. In this paper, a pooled model proposed by Oaxaca & Ransom (1994) is used to estimate β^*. Using Oaxaca & Ransom (1994) method, equation (4) can be demonstrated as below

$$\begin{aligned} D = \{E(X_M) - E(X_F)\}' \{ W\beta_M + (I - W)\beta_F\} + \\ \{(I - W)' E(X_M) + W' E(X_F)\}' (\beta_M - \beta_F) \end{aligned}$$

where $\widehat{W} = \Omega = (X_M' X_M + X_F' X_F)^{-1} X_M' X_M$

(7)

W is a matrix of relative weights given to the coefficients of male,

I is the identity matrix, and

X is the matrix of observations

Table 7 represents the result of the decomposition analysis on the pooled sample. The gender log wage differential is 0.521. The geometric mean of wages is 300.541 KRW and 178.424 KRW for men and women, respectively, which is equivalent to a difference of 68.4%. If the levels of women's endowments are improved to those of men, wages of women will rise by 19.2%, but the adjusted wage gap of 41.3% shows a greater portion of the wage gap cannot be explained by differences in characteristics. As noted earlier, this unexplained part could be attributed to either discrimination or missing variables. Positive contribution of a variable in the unexplained part suggests that male wage-earners with a certain level of endowment are rewarded higher compared to female workers with a similar degree of characteristics. Subsequently, negative contribution implies the opposite, which narrows the wage disparities.

Table 7. Blinder-Oaxaca Decomposition Result

	lnwage	Exp(b)	%
Female	5.184*** (0.012)	178.424	
Male	5.706*** (0.010)	300.541	
Difference	0.521*** (0.015)	1.684	100
Explained	0.176*** (0.011)	1.192	33.723
Unexplained	0.346*** (0.013)	1.413	66.277
	explained	%	unexplained
Age	0.002 (0.001)	0.294	0.235*** (0.066)
Edu	0.028*** (0.004)	5.354	0.086 (0.087)
Exp	0.056*** (0.008)	10.733	-0.022 (0.029)

Exp2	-0.008 (0.005)	-1.486	-0.021 (0.013)
Mar	0.005*** (0.002)	1.054	0.145*** (0.018)
Area	-0.001 (0.001)	-0.195	0.021 (0.012)
Ft	0.038*** (0.005)	7.297	-0.074* (0.032)
Job	-0.003* (0.001)	-0.566	-0.002 (0.009)
Union	0.007** (0.002)	1.264	-0.007 (0.005)
1-30 employees	0.017*** (0.002)	3.238	-0.003 (0.010)
30-300 employees	-0.001* (0.001)	-0.266	-0.013** (0.005)
300 and above employees	0.012*** (0.002)	2.331	0.011* (0.005)
Primary	-0.000 (0.000)	-0.043	0.001 (0.001)
Secondary	0.018** (0.006)	3.489	0.003 (0.012)
Tertiary	0.007 (0.006)	1.257	-0.035 (0.026)
Quaternary	-0.000 (0.001)	-0.033	-0.006 (0.008)
Constant			0.027 (0.128)
N	4298		

Robust standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

It is shown that out of the gender wage difference of 68.4%, 33.723 percentage points can be attributed to distinctness in the covariates whereas 66.277 percentage points can be credited to discrimination or the potential effects of differences in omitted variables. The contribution of occupation to the sex pay gap is −0.5666%. This means women in the sample tend to have high-skilled jobs compared to men in the data set. The level of education and the years of experience accounts for 0.028 and 0.056 of the total differences separately together equal to over 15% of the gender pay gap. Residing in the Greater

Seoul would decrease the gender wage gap slightly, albeit its statistical insignificance. Working in the secondary industry explains 3.489% of the wage difference which is consistent with the finding of Fields & Wolff (1995) that the gender wage gap of 15-19% could be partly explained by segregation of women in low paying industries by using the U.S. CPS data of March 1988. With the 1999 KLIPS data, Heo (2003) also found similar evidence in the analysis of the gender pay differences in industry level in Korea.

In the unexplained part, it is notable that age and marital status are statistically significant at 1%. Positive signs on age and marital status in the unexplained component mean they are unfavourable to women: older or married women would experience wider income differentials. This may imply discrimination against women by age and marital status; latter sometimes termed as a marriage penalty on women. In much economic literature on income, age is regarded as “a proxy for potential work experience” (Lee et al., 2008) and represent seniority. Hence, it is expected to have a positive relationship with wage as well as labour force participation rate up to a certain age. However, as observed in the estimation of the regression model, the age variable has opposite effects on wages by gender which are statistically significant only on male. Abe (2010) found the similar result in Japan that the gender income gap widens with age.

In addition, marital status has the opposite effect on wages by gender, showing statistically significant only on males in the regression analysis followed by statistical significance in both explained and unexplained gender wage differentials. The wider gender wage gap for married workers are observed in other countries (Blau and Kahn, 1992; Polachek, 2004; Hughes & Maurer-Fazi, 2002). Hughes & Maurer-Fazi (2002), in their paper exploring the effects of marriage, education and occupation on the gender wage gap in China, found that the gender income gap is larger for married workers compared to single workers and women experience a marriage penalty regarding larger unexplained wage gaps compared to married men holding similar level of productive characteristics. Blau and Kahn (1992) illustrated same findings that female/male earning ratios are lower for married workers compared to single workers in Western countries, i.e., Germany, United Kingdom, U.S., Austria, Switzerland, Sweden, Norway and Australia. Using the U.S. Census Data of 1990, Polachek (2004) found diminishing gender wage gap with age but distinguishable wage profiles for married men and women: “Married men have far higher and steeper earnings profiles than married women”.

The female age and marriage penalty could be associated with the labour force participation rate for women of which the curve is M-shaped or the twin-peaked that reflects intermittent of work for female. As pointed by Tanaka & Muzones (2016) and Lee et al. (2008), Korean women tend to leave the workforce after childbirth and during the childbearing years while men are likely to remain in the labour force even though young men and women enter into the labour force at about the same age. With wide-spread seniority-based wages and promotion system generally adopted by establishments in Korea, this could be a reason behind the positive contribution of marital status and age on the adjusted gender wage gap.

3. Matching Study

Selection bias is one of the major problems in a non-experimental setting. The observational data on wages are not available for the unemployed, which could make the employed as selective groups. In fact, it is using the data compiled from the subsample of wage earners selected non-randomly. When analysing the gender wage gap, selection bias becomes more problematic if the factors affecting selection into the labour force were different by sex. When female employment rate is low, women who are likely to receive high salaries would enter the labour market. Boll & Lagemann (2018) showed the evidence of a strong employment selection of women in European countries with low female labour force participation rate below 65% utilising the Structure of Earnings Survey (EU-SES) for 2014. This could produce biased coefficient estimates in the wage equation especially for the characteristics affecting the probability of working and productivity (Goraus et al., 2017).

The heterogeneity of the sample is another problem. Ñopo (2008) argued that the result from the Blinder-Oaxaca decomposition could be misleading due to “differences in the supports", i.e., the differences in the distributions of characteristics held by men and women. For instance, a single individual holding a bachelor's degree and working in a firm with more than 300 employees may be found in one gender group but not in the other. With lack of comparable samples that share similar characteristics, the Blinder-Oaxaca decomposition analysis is based on an "out-of-support assumption" which in turn results in the over- or under- estimation of the gender wage gap caused from differences in the premiums (Ñopo, 2008).

The matching approach could address the problems of selection bias and the differences in common supports. The basic idea of matching approrach is comparing like with like to analyse the outcome so that it is possible to estimate the pure effect of the treatment/assignment. It aims to choose samples of the untreated and the treated who are alike in observed covariates. However, conditioning on many explanatory variables is likely to produce poor matching, so-called the "curse of dimensionality" but the propensity score suggested by Rosenbaum & Rubin (1983) could be used to deal with this problem. Rosenbaum & Rubin (1983) defined that "The propensity score is the conditional probability of assignment to a particular treatment given a vector of observed covariates" and suggested it can be applied to produce matched samples.

$$\text{Propensity Score } P(X) \equiv P(F=1|X) = E(F|X)$$

Assuming F is a treatment variable which equals 1 for female and 0 for male and the log wage is be defined as Y , the treatment effect is

$$D = Y(1) - Y(0).$$

Then the average treatment effect (ATE) becomes:

$$D_{ATE} = E(D) = E[Y(1) - Y(0)]$$

and the average treatment effect of the treated (ATT), in this case the average effect of being women on wages can be written as:

$$D_{ATT} = E(D|\,F=1) = E[\,Y(1)\,|F=1] - E[\,Y(0)\,|\,F=1]$$

(8)

The problem of estimating ATT in equation (8) is that the counterfactual situation, $E[\,Y(0)\,|\,F=1]$, the expected log wage of woman if she was a man cannot be estimated. Rosenbaum & Rubin (1983) proved that under the assumptions of unconfoundedness $Y(0),\ Y(1) \perp D|X$ and overlap $0 < P(F=1|X) < 1$, the average treatment effect, D_{ATT}, can be defined for all explanatory variables X. The first one, unconfoundedness or conditional independence assumptions (CIA), infers that "systematic differences in outcomes between treated and comparison individuals with the same values for covariates are attributable to treatment" (Caliendo & Kopeinig, 2008). The second one, overlap conditions, means that given a set of control variables X, an individual can be men (control) or women (treatment) with a positive probability. (Heckman et al., 1999). If the assumptions of conditional independence and overlap hold, the equation (7) becomes

$$D_{ATT}^{PSM} = D_{P(X)|F=1}\{E[\,Y(1)|F=1, P(X)] - E[\,Y(0)|F=0, P(X)]\}$$

(8)

In Caliendo and Kopeinig's (2008) term, "the PSM estimator is simply the mean difference in outcomes over the common support, appropriately weighted by the propensity score distribution" of the treatment group.

Only problem of using the propensity score matching in analysis of the gender wage gap is that CIA is not likely to be met when being female is perceived as "treatment". Yet, in the context of wage analysis, the assumption of unconfoundness may indicate "after controlling for X, there are no unobserved characteristics which are productivity relevant to explain wages (Djurdjevic, D. & Radyakin, S., 2005)". As a result, if incomes are found to be different by gender, they can be devoted to discrimination. Or alternatively, matching on propensity score could still be valid referring Frölich (2007) who justified the use of matching on propensity score outside of traditional treatment evaluation, for instance, "as a nonparametric extension of the Blinder (1973) Oaxaca (1973) decomposition". He argued that it did not hinge on CIA and proved the consistency of propensity score matching estimator based on finite mean assumptions given the justification of matching on covariates. In his work, propensity score matching was applied to analyse the impact of the subject of degree on gender pay differences in England (Frölich, 2007).

The Propensity score can be estimated by probit or logit model. After the propensity score is calculated based on various covariates, the neighbourhood for each treated sample is to be chosen based on a matching algorithm that decides the weight and the technique for the matching partner selection (Caliendo & Kopeinig, 2008). In this study, the propensity score is estimated with the probit model, and the result can be found in table 11 in Appendix.

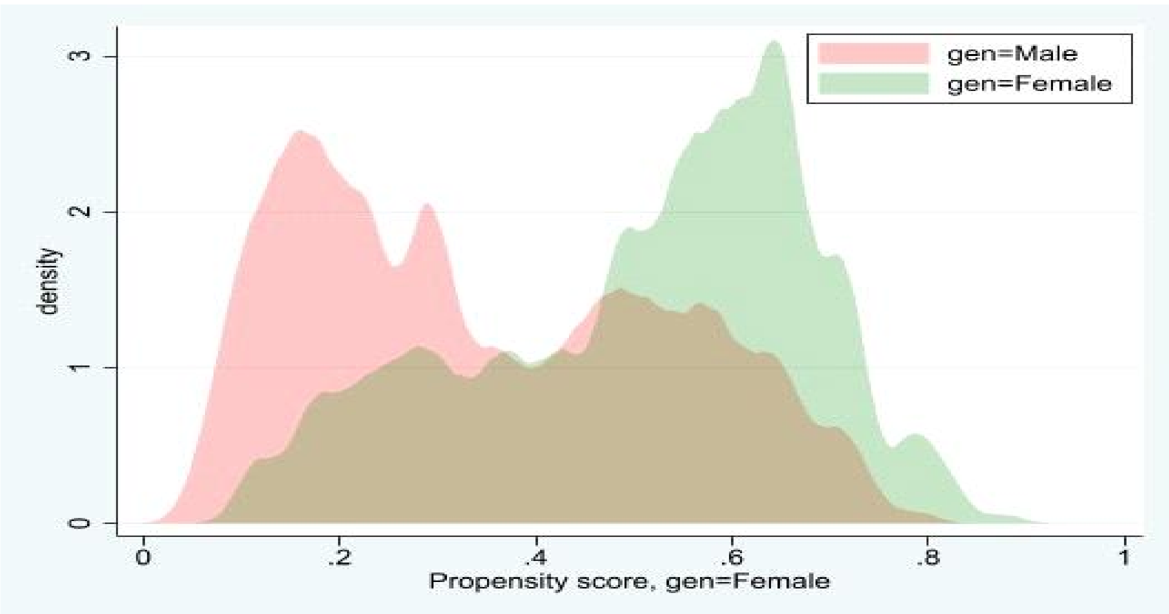

Figure 2. Overlap condition based on the estimated propensity score

Figure 2 is to show whether the overlap assumption could be satisfied. It displays the estimated density of the predicted probabilities that a woman is a woman and the estimated density of the predicted probabilities that a man is a woman. There is no evidence that the overlap condition is violated because none of the plots is gathered near 0 or 1.

Technical features of the different matching algorithm are beyond the scope of this paper, but the basic concepts are described briefly referring Caliendo & Kopeinig (2008). K-Nearest Neighbourhood (NN) matching selects a K number of matching partner that has the closest propensity score. K can be chosen, and the one from the control group can be selected multiple times with replacement. When matching with replacement 1:K is allowed, a trade-off between bias and variance would rise. Means variance would decrease because more information was used to build the counterfactual but due to the possibility of poor matching, the bias would rise.

The problem of NN matching is that it might produce inadequate matches. To avoid matching on an irrelevant neighbour, Cochran and Rubin (1973) proposed Caliper or Radius matching. A "caliper" which is a maximum propensity score distance for potential matches, therefore, it can ensure a better quality of matching. Yet the caliper should be chosen by a researcher a priori.

Kernal matching uses the weighted averages of almost all members in the control group to construct the counterfactual outcome, hence the loss of information is minimised. The closer PSM distance potential matches have, the bigger weights are given. Kernel function and bandwidth parameter have be to decided but with the decision of low bandwidth a small variance can be only achieved in exchange of the unbiasedness in the estimator.

In this paper, 1:1 nearest neighbourhood matching with replacement is employed on the original sample to create matches that are close enough in terms of the propensity score. A caliper of 0.1 is set to limit the risk of poor matches of men to women but to keep the as much information as possible. As a result of matching on the estimated propensity score, 1,786 males are matched to 1,786 females.

Figure 3 exhibits the distributions of the propensity scores before and after matching. It is clear that the densities of them are more homogeneous after matching and there is an apparent overlap of the distributions. The change in the distribution of mean in variables

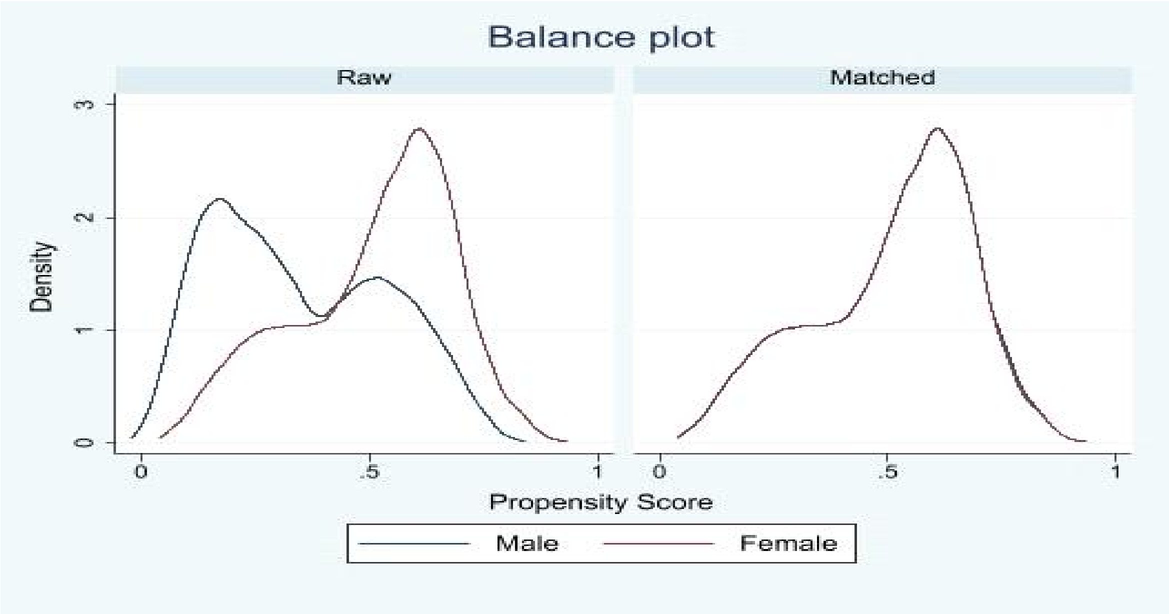

Figure 3. Balance plot Before and After matching

before and after matching is also presented in table 8. As a result of matching, the number of samples are reduced but only sample of men and women who share similar characteristics chosen for the comparison. Before matching, male works have longer years of experience, however, matched men have shorter years of experience. A big reduction in the mean of the secondary industry for men indicates some samples of men working in the secondary industry are discarded because they could not find matching partners from samples of women. Bias adjustment is the smallest for the residential area because over 50% of men and women reside in the Greater Seoul as shown in table 4. Table 12 in Appendix shows the standardised differences and variance ratio for the original data and the matched sample. The standardised differences close to zero, and the variance ratios close to one indicates propensity score matching balances the control variables. However, the inference is informal without standard errors.

Table 8. Distribution of Mean Before and After Matching

	Before		After			
	Female	Male	Female	Male		
	N=1,786	N=2,512	N=1,786	N=1,786	t	p>|t|
Age	41.012	41.887	41.012	40.800	90.7	0.552

Edu	13.368	14.044	13.368	13.269	59.2	0.220
Exp	5.1321	7.864	5.132	5.211	77.8	0.695
Exp2	60.447	121.19	60.447	64.624	79.3	0.382
Mar	0.633	0.691	0.633	0.612	67.7	0.202
Area	0.534	0.506	0.534	0.558	15.9	0.139
Ft	0.745	0.859	0.745	0.737	72.3	0.620
Job	0.304	0.269	0.304	0.292	82.8	0.421
Union	0.091	0.202	0.092	0.089	70.0	0.771
1-30 employees	0.585	0.428	0.585	0.607	64.4	0.173
30-300 employees	0.243	0.306	0.243	0.244	42.9	0.969
Secondary	0.185	0.488	0.185	0.199	72.5	0.270
Tertiary	0.653	0.372	0.653	0.632	72.8	0.185
Quaternary	0.158	0.130	0.158	0.167		0.441

The average treatment effect on the treated and the average treatment effect are calculated on matched samples as in table 9. The natural logarithm of the wage for unmatched female and male are 5.184 and 5.703, respectively, which are equivalent to the previous results. The difference has a negative sign as the treatment variable F is set to 1 for female, indicating the average wage of females is lower. The average wage difference between men and women in the matched sample represented by ATE equals −0.324 and ATT is −0.320 which means for a given level of wage-relevant characteristics women on average would receive 32% lower monthly income.

Table 9. Matching Result

Variable	Sample	Female	Male	Difference	AI Robust S.E.	[95% Conf. Interval]	
lnwage	unmatched	5.184	5.703	-0.521			
	ATT	5.184	5.504	-0.320***	0.022	-0.363	-0.275
	ATE			-0.324***	0.022	-0.367	-0.282

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

From the regression model with sex dummy variable represented by equation (2), the estimated coefficient of β_7 is equivalent to ATT as it measures the size of the gender wage gap after controlling for the effect of other covariates on wages (Goraus et al., 2017). As recognised by several economists (Barsky et al., 2002; Black et al., 2006; Kline, 2011; Melly, 2006) the unexplained part of the Blinder-Oaxaca decomposition can be evaluated as ATT as well. Kline (2001) proved that "The regression based Oaxaca-Blinder estimator of counterfactual means is equivalent to a propensity score reweighting estimator modeling the odds of treatment as a linear function of the covariates". Melly (2006) also revealed "the Oaxaca/Blinder decomposition estimates the average treatment effect on the treated", showing the effects of coefficients, $\overline{X_M}'(\widehat{\beta_M}-\widehat{\beta_F})$ or $\overline{X_F}'(\widehat{\beta_M}-\widehat{\beta_F})$, in the B-O decomposition can be written as $E[Y(1)|F=1]-E[Y(0)|F=1]$ under the assumption of unconfoundedness and traditional Oaxaca (1973) and Blinder (1973) assumption, "the expected value of Y conditionally on X is a linear function of X". Hence ATT can be interpreted as the unexplained part in the B-O decomposition, $\left\{E(X_M)'\left(\beta_M-\beta^*\right)+E(X_F)'\left(\beta^*-\beta_F\right)\right\}$. Contrast to the B-O decomposition, the counterfactual wage for female, $E[Y(0)\,|\,F=1]$, is not estimated when propensity score matching is applied but set as the average wage of matched men (Pacheco et al., 2017). Because males in the matched sample have similar characteristics as females, their average wages can be regarded as the wage of females as if they were men.

Table 10. Gender Wage Gap from Different Methods

	OLS	B-O decomposition	PSM
Adjusted Wage gap	-0.337	0.346	-0.320

The result of the OLS model, the B-O decomposition and propensity score matching is summurised in table 10. Regardless of the specifications, the size of the adjusted wage gap is substantial showing women receive, on average, lower monthly income compared to men by from 41% to 32%. Even though the size of gender wage differentials is the smallest when estimated with the semi-parametric method based on the common support, the gender wage differentials remain fairly steady. In addition, similarity of these estimates might indicate that selection into employment has minimal effect on the gender wage gap. Women may find it difficult to join the labour force, but once they are employed, these are not related to the gender wage difference (Goraus et al., 2017).

V. Conclusion

In this paper, the adjusted gender wage gap in Korea is investigated with the 2018 KLIPS data employing parametric and semi-parametric approach. First, the effect of social-economic characteristics including gender dummy on wage is analysed with the regression model. The OLS estimates on pooled samples show being female decreases the log average monthly wage by 33.7% whereas the years of educations and experiences, the regularity of work, having a high-skill job and working in the larger workplace increase it. Age and marriage have positive contributions on wages of the pooled sample, but this is because their positive relationships with the men's earnings surpass their negative relationships with those of females.

Second, Oaxaca & Ransom (1994) decomposition analysis is used to break the raw wage gap into one that is caused by differences in the endowments, and the other that cannot be attributed to those differences. The result shows that the share of unexplained part is larger than the explained part indicating an issue of omitted variables or the existence of discrimination against women. The twin-peaked nature of the female labour force participation rate and the seniority-based pay system in Korea could be the reasons behind positive contributions of age and marital status on the unexplained differences.

Despite the advantages of Blinder-Oaxaca decomposition, it has weaknesses if there is the heterogeneity in the socio-economic characteristics by gender or the sample selection bias. To address these issues, the propensity score matching method is utilised and its result is compared with the estimates from the traditional parametric method described earlier. Although the adjusted wage gap estimated based on comparing with the comparable by the propensity score shows the lowest gap, 32%, the differences in the size of the gender wage gap are not distinctive. This implies the policy to narrow the gender gap in Korea should be directed not only to increase the level of the human endowment for women but also to reduce the size of unexplained gender wage gap derived by the social norms and the attitude to the gender role.

Yet it should be noted that matching on the propensity score also has its drawbacks. One of them is that the results are meaningful only within the common support (Goraus et al., 2017). The application of the results from this study could be limited due to the small number of samples. The second disadvantage is that the matching may not address selection bias completely, even though this issue has not been discussed closely in this paper. Men and Women might have different unobserved characteristics which make a certain group more likely to work or to be employed. As Becker's (1971) discrimination theory suggests, if women are required to have more market-favourable skills to be employed, the unexplained component of the wage gap would be underestimated. Therefore the quality and size of the sample are important. Including variables such as the number of children, the non-labour income and so on could address the selection bias and might be able to reduce the size of the adjusted gap, which is left for future research.

To be noted, this study is only complementary to the existing literature on the gender wage gap in Korea. Still its contribution lies in the comparison of like with like in terms of gender pay differences using the most recent available data. This work also provides empirical evidence of gender wage disparities among those who have similar characteristics.

Reference

Abadie, A., & Imbens, G. (2006). Large Sample Properties of Matching Estimators for Average Treatment Effects. *Econometrica*, 74(1), 235-267. Retrieved June 6, 2020, from www.jstor.org/stable/3598929.

Abe, Y. (2010). Equal Employment Opportunity Law and the gender wage gap in Japan: A cohort analysis. *Journal of Asian Economics*, 21(2), 142-155, https://doi.org/10.1016/j.asieco.2009.12.003.

Ahn, T. (2012). Gender wage gap across the wage distribution in Korea. *Applied Economics*, 14(1), 127-149.

Ahn, J. K. & Oh, S.M. (2017). A Study on the Gender Wage Gap in Youth. *KEIS Working Paper*, Korea Employment Information Service (KEIS), May 2017.

Austin, P. C. (2014). A comparison of 12 algorithms for matching on the propensity score. *Statistics in medicine*, 33(6), 1057-1069.

Barsky, R., Bound, J., Charles, K., & Lupton, J. (2002). Accounting for the Black-White Wealth Gap: A Nonparametric Approach. *Journal of the American Statistical Association*, 97(459), 663-673. Retrieved July 5, 2020, from www.jstor.org/stable/3085702.

Becker, G. (1971). *The Economics of Discrimination*. Chicago: University of Chicago Press.

Becker, G. (1985). Human capital, Effort and the sexual Division Labour. *Journal of Labour Economics*, 3(5), 33-58.

Bergmann, B. (1974). Occupational Segregation, Wages and Profits When Employers Discriminate by Race or Sex. *Eastern Economic Journal*, 1(2), 103-110. Retrieved June 29, 2020, from www.jstor.org/stable/40315472.

Blau, F., & Kahn, L. (2000). Gender Differences in Pay. *The Journal of Economic Perspectives*, 14(4), 75-99. Retrieved June 28, 2020, from www.jstor.org/stable/2647076.

Blinder, A. S. (1973). Wage Discrimination: Reduced Form and Structural Estimates. *Journal of Human Resources*, 8(4), 436–455, https://doi.org/10.2307/144855.

Boll, C., & Lagemann, A. (2018). Gender pay gap in EU countries based on SES (2014).

Luxembourg, Publication Office of the European Union. doi, 10, 978935.

Caliendo, M., & Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of economic surveys*, 22(1), 31-72.

Cochran, W., & Rubin, D. (1973). Controlling Bias in Observational Studies: A Review. Sankhyā: *The Indian Journal of Statistics*, Series A (1961-2002), 35(4), 417-446. Retrieved June 6, 2020, from www.jstor.org/stable/25049893.

Cotton, J. (1988). On the Decomposition of Wage Differentials. *The Review of Economics and Statistics*, 2, 236.

Daly, A., Meng, X., Kawaguchi, A., & Mumford, K. (2006). The gender wage gap in four countries. *Economic Record*, 82(257), 165-176.

Djurdjevic, D. & Radyakin, S. (2005). Decomposition of the gender wage gap using matching: an application for Switzerland, *Darmstadt Discussion Papers in Economics*, No. 155, Technische Universität Darmstadt, Department of Law and Economics, Darmstadt.

Fields, J., and Wolff, E. (1995). Interindustry Wage Differentials and the Gender Wage Gap. *Industrial and Labor Relations Review*, 49(1), 105-120. doi:10.2307/2524915.

Frölich, M. (2007). Propensity score matching without conditional independence assumption —with an application to the gender wage gap in the United Kingdom. *The Econometrics Journal*, 10(2), 359-407. Retrieved July 3, 2020, from www.jstor.org/stable/23116034.

Goraus, K., Tyrowicz, J., & Van der Velde, L. (2017). Which gender wage gap estimates to trust? A comparative analysis. *Review of Income and Wealth*, 63(1), 118-146.

Grybaite, V. (2006) Analysis of theoretical approaches to gender pay gap. *Journal of Business Economics and Management*, 7(2), 85-91.

Jonathan Guryan & Kerwin Kofi Charles. (2013). Taste-based or Statistical Discrimination: The Economics of Discrimination Returns to its Roots. *Economic Journal*, Royal Economic Society, 123(11), 417-432.

Heo, S. (2003). An Empirical Analysis of Gender Wage Differentials in Industries and Occupations. *Applied Economics*, 5(3), 57-74.

Heckman, J., Ichimura, H. & Todd, P. (1998) Matching As an Econometric Evaluation

Estimator. *Review of Economic Studies*, 65(2), 261-294.

Hughes, J. & Maurer‑Fazio, M. (2002), Effects of Marriage, Education and Occupation on the Female / Male Wage Gap in China. *Pacific Economic Review*, 7: 137-156. doi:10.1111/1468-0106.00156.

Jang, J. (2020). Effect of Marriage and Infant Child on Gender Wage Gap. *Ewha Journal of Gender and Law*, 12(1), 179-212.

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

Kang, J., & Kim, G. (2014). The Relationship between Gender Wage Gap and Occupational Segregation. *Journal of Labour Economics*, 37(4), 113-141.

Kim, T. (2013). The Decomposition of Wage Gap due to Gender and Employment Pattern. *The Women's Studies*, 84(1), 31-61.

Lee, B. S., Jang, S., & Sarkar, J. (2008). Women's labor force participation and marriage: The case of Korea. *Journal of Asian Economics*, 19(2), 138-154.

Lemieux, T. (2006). The "Mincer equation" thirty years after schooling, experience, and earnings. In *Jacob Mincer a pioneer of modern labor economics* (127-145). Springer, Boston, MA.

Meara, K., Pastore, F. & Webster, A. (2019). The gender pay gap in the USA: a matching study. *Journal of Population Economics*. 10.1007/s00148-019-00743-8.

Mincer, J. (1974). *Schooling, Experience and Earnings*. New York: National Bureau of Economic Research.

Mincer, J. & Polachek, S. (1974). Family investment in human capital: Earning of Women. *Journal of Political Economy*, 82(2), 96-108.

Oaxaca, R. (1973). Male-Female Wage Differentials in Urban Labor Markets. *International Economic Review*, 14(3), 693-709. doi:10.2307/2525981.

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

OECD (2020), Gender wage gap (indicator). doi: 10.1787/7cee77aa-en. Accessed on 28 June 2020.

Pacheco, G., Li, C., & Cochrane, B. (2017). Empirical evidence of the gender pay gap in

New Zealand. *Wellington: NZ Work Research Institute, Auckland University of Technology.*

Polachek, S. (2004). "How the Human Capital Model Explains Why the Gender Wage Gap Narrowed," *IZA Discussion Papers 1102*, Institute of Labor Economics (IZA).

Rosenbaum, P., & Rubin, D. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika*, 70(1), 41-55. doi:10.2307/2335942.

Sianesi, B. (2001). Implementing propensity score matching estimators with Stata. In UK Stata Users Group, VII Meeting.

Shin, K. (2011) The Gender Pay Gap in South Korea : Difference and Discrimination. *Korean Journal of Sociology,* 45(4), 2011.8, 97-127.

Sun, H. & Kim, G. (2015). Human Capital and Glass Ceiling: Quantile Regression Decomposition of Gender Pay Gap in Korean Labor Market. *Global Journal of Human-Social Science Research*, 15.

Tanaka, S., & Muzones, M. (2016). Female labor force participation in Asia: Key trends, constraints, and opportunities. *ADB Brief*, No. 71. Manila: Asian Development Bank.

Yoo, J. (2017). An Analysis of the Gender Wage Gap in Youth Generation. *Journal of Korean Women's Studies*, 33(1), 107-155.

Appendix.

Table 11. Propensity score estimation using probit regression

Treatment= Female	Coef.	Std.Err.	z	P>|z|	[95% Conf.	Interval]
Age	-0.004	0.003	-1.47	0.140	-0.009	0.001
Edu	-0.106	0.010	-10.27	0.000	-0.126	-0.086
Exp	-0.031	0.009	-3.64	0.000	-0.048	-0.014
Exp2	0.000	0.000	0.71	0.478	0.000	0.001
Mar	0.060	0.049	1.22	0.221	-0.036	0.156
Area	0.030	0.041	0.71	0.476	-0.052	0.111
Ft	-0.170	0.055	-3.08	0.002	-0.278	-0.062
Job	0.154	0.052	2.95	0.003	0.051	0.256
Union	-0.255	0.072	-3.55	0.000	-0.396	-0.114
1-30 employees	-0.020	0.065	-0.31	0.760	-0.147	0.107
30-300 employees	-0.062	0.066	-0.94	0.346	-0.190	0.067
Secondary	0.219	0.279	0.79	0.432	-0.328	0.766
Tertiary	1.115	0.278	4.01	0.000	0.569	1.660
Quaternary	1.051	0.284	3.7	0.000	0.494	1.608
Constant	0.875	0.333	2.63	0.009	0.222	1.528

Log likelihood = -2541.601 Prob > chi2 = 0.000 Pseudo R2 = 0.1289

Table 12. Standardised Differences and Variance Ratio Comparison : Before and After Matching

	Standardised Diff		Variance Ratio	
	Before	After	Before	After
Age	-0.085	0.020	1.319	1.130
Edu	-0.282	0.041	0.993	0.991
Exp	-0.400	-0.013	0.575	0.910
Exp2	-0.341	-0.029	0.407	0.817
Mar	-0.124	0.043	1.089	0.979
Area	0.052	-0.049	0.996	1.009
Ft	-0.289	0.017	1.567	0.982
Job	0.077	0.027	1.076	1.024
Union	-0.316	0.010	0.517	1.028
1-30 employees	0.318	-0.046	0.992	1.018
30-300 employees	-0.142	-0.001	0.866	0.998

Secondary	-0.678	-0.037	0.603	0.944
Tertiary	0.586	0.044	0.969	0.974
Quaternary	0.079	-0.026	1.174	0.954