Male-female wage differentials in the San Francisco Bay Area

Ganji, Niloofar

California State University-East Bay

20 June 2019

Male-Female Wage Differentials in The San Francisco Bay Area

Niloofar Ganji†

†Department of Economics, California State University-East Bay, 25800 Carlos Bee Blvd, Hayward, CA 94542

†E-mail: nganji@horizon-csueastbay.edu

Abstract

Despite all efforts in the past few decades, the pace of male-female wage convergence was very slow. San Francisco Bay Area as an advanced, culturally diverse and a pioneer socioeconomically metropolitan suffers from male-female wage discrimination. It is the purpose of this study to estimate the average extent of discrimination against female workers in the Bay Area and to provide a quantitative assessment of the sources of wage inequality using Blinder-Oaxaca decomposition. The results indicate that influencing factors such as education, marital status, and work experience are rewarded differently between men and women. Additionally, the major portion of the wage-gap remains unexplained with human capital characteristics.

Keywords:

Gender wage-gap, Blinder-Oaxaca decomposition, Human capital characteristics, Twofold decomposition

1) Introduction

The gender wage-gap has been intensively investigated for a number of decades but also remains an area of active and innovative research. Despite all policies enforced by lawmakers and corrective actions taken by corporations in the past decades, the gender wage differentials still remain high (Blau and Kahn 2017).

This article provides new empirical estimates describing the extent of and trends in the gender wage-gap and their potential explanations. The primary focus will be on the San Francisco Bay Area as a good example to study the gender wage-gap extent in an advanced metropolitan area.

The focus on the Bay Area is, in part, designed to make the task more specific and manageable, as there have been numerous research studies on this topic across the United States and many other countries. Nonetheless, the author believes much of what has been learned for the Bay Area is applicable to other parts of the United States and the world, particularly other economically advanced metropolitan areas with similar characteristics.

Policymakers have the long-lasting concern opposing wage discrimination against women. Therefore, the assessment of such policies cannot be accomplished without a precise measure of the unequal treatment of male and female workers. Economists commonly define wage discrimination by comparing wages for equally productive workers (Becker 1964).

The wage-gap is usually estimated by a difference in wage on those of human capital characteristics that reflect productivity potential (Kunze 2005). In this case, the logarithmic wages (ln (wage)) are regressed on measures for individual work histories, education and other background variables (Jann 2008). In addition, the wage gap is decomposed into “explained” and “unexplained” portions based on human capital characteristics. The unexplained part of the wage-gap is referred to as an estimate of wage discrimination (Kunze 2008).

The next section of this article briefly discusses the data treatments and the Blinder-Oaxaca decomposition method used to study the gender wage-gap in the San Francisco Bay Area.

2) Materials and Methods

The standard procedure to investigate differences in wages is the one developed by Blinder (1973) and Oaxaca (1973) which allows that productive characteristics of men and women are rewarded differently. The Blinder-Oaxaca decomposition is a statistical method that decomposes differences in mean outcomes across two groups into a part that is due to group differences in the levels of explanatory variables and a part that is due to differential magnitudes of regression coefficients (Marek Hlavac 2018).

Here the two groups are labeled as Group M (for males) and Group F (for females). The mean outcome difference to be explained ($\Delta\bar{Y}$) is simply the difference of the mean outcomes for observations in Group M and Group F, denoted as $\bar{Y}_M$ and $\bar{Y}_F$, respectively:

$$\Delta\bar{Y} = \bar{Y}_M - \bar{Y}_F \tag{1}$$

The above equation can be decomposed using either threefold decomposition or twofold decomposition. Here, twofold Blinder-Oaxaca decomposition is utilized for estimation (in this case, threefold decomposition gives similar results). The twofold approach decomposes the mean outcome difference with respect to a vector of reference coefficients $\hat{\beta}_R$. As Equation 2 shows, the twofold decomposition divides the difference in mean outcomes into a portion that is explained by cross-group differences in the explanatory variables, and a part that remains unexplained by these differences.

The unexplained portion of the mean outcome gap has often been attributed to discrimination, but may also result from the influence of unobserved variables (Kunze 2008). The unexplained part can be further broken down into "unexplained male" and "unexplained female".

$$\Delta\bar{Y} = \underbrace{(\bar{X}_M - \bar{X}_F)'\hat{\beta}_R}_{\text{explained}} + \underbrace{\underbrace{\bar{X}'_M(\hat{\beta}_M - \hat{\beta}_R)}_{\text{unexplained M}} + \underbrace{\bar{X}'_F(\hat{\beta}_R - \hat{\beta}_F)}_{\text{unexplained F}}}_{\text{unexplained}} \tag{2}$$

There are different ways of calculating $\hat{\beta}_R$. In this study, the method suggested by Neumark (1988) is utilized and reference coefficients ($\hat{\beta}_R$) are taken from a pooled regression that does not include the gender indicator variable.

The data for the present study are collected from the IPUMS CPS database[1] for the individuals who were full-time employed during the period of three years (2016-2018) and worked in the San Francisco Bay Area. For each individual information about age, gender, race, marital status, years of education, and hourly wage are extracted. The hourly wage has been used as it is more precise compared to weekly or monthly wages (see Figure 1). In order to account for inflation, all the wages are adjusted based on their 2018 dollar value. An important parameter that explains a significant portion of the wage difference between workers is work history which is not available from most databases. Therefore, the following equation was used to estimate the work experience for individual i.

$$X_{experience,i} = X_{age,i} - X_{education,i} - 6 \tag{3}$$

Where, $X_{experience,i}$ is years of work experience for individual i, $X_{age,i}$ is the age of individual i, $X_{education,i}$ is years of schooling for individual i; and 6 is the age of start of

[1] www.cps.ipums.org

school. Based on the trend in data, the binary variable of "work experience" was defined as 1 for $X_{experience,i} \geq 10$ and 0 for $X_{experience,i} < 10$.

For modeling using R software, the variable "married" has been selected as an omitted variable giving the least some of the error (SE). The sample population consists of 42% female, 58% married, 41% white and 80% with a college education.

3) Results and Discussion

The linear regression model includes covariates that account for the workers' marital status, race, work experience and years of a college education. Worker's age is not included to avoid collinearity, as the age is used to estimate the work experience. log-wage is a dependent variable that denotes ln(hourly-wage).

The coefficients of group regressions are listed in Table 1 and all are statistically significant except race (white). From the coefficients, it can be seen that individuals with less human capital (not-married individuals, less work experience and those without a college education) tend to earn less. Being married has the largest influence on the wage for males and work experience for females. Years of college education indicates that women with no college education will face larger wage discrimination.

Table 1. Linear Regression Results

	Dependent variable:		
	log-wage		
	male	female	pool
	(1)	(2)	(3)
Married	0.218***	0.145***	0.200***
	(0.032)	(0.032)	(0.024)
White	0.028	0.007	0.026
	(0.029)	(0.029)	(0.022)
Work_Experience	0.220***	0.358***	0.284***
	(0.038)	(0.038)	(0.028)
Yrs_College_Edu	0.127***	0.147***	0.133***
	(0.005)	(0.005)	(0.004)
Constant	2.822***	2.520***	2.687***
	(0.041)	(0.041)	(0.031)
Observations	1,839	1,320	3,159
R2	0.285	0.314	0.291
Adjusted R2	0.284	0.312	0.290

Note: *p<0.1; **p<0.05; ***p<0.01

The mean log-wages is 3.559 for the male workers and 3.370 for female workers, leaving the difference of approximately 0.189 (17.3% of the wage gap) to be explained by the Blinder-Oaxaca decomposition. Furthermore, the value of predictors' mean difference ($\Delta\bar{X} = \bar{X}_M - \bar{X}_F$) shows that there is not a significant difference between male and female in terms of marital status and work experience ($\Delta\bar{X} = 0.083$ and 0.037, respectively). Actually, the years of college education indicate that the mean of this predictor for females is slightly higher than males ($\Delta\bar{X} =$ -0.154).

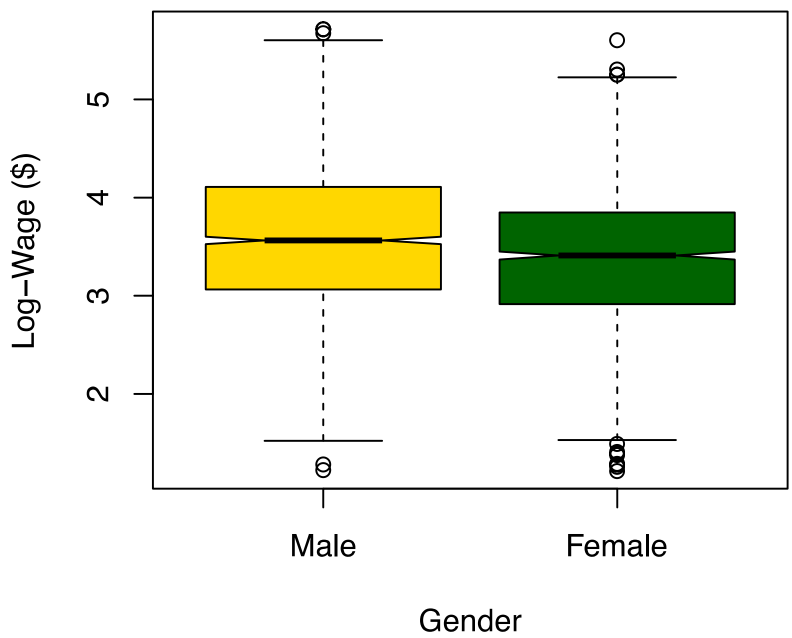

Figure 1. Boxplot for comparison of the female vs. male hourly wages

The wage-gap is divided into "explained" and "unexplained" parts (see Figure 2). As the results show, the major part of the wage-gap falls within the unexplained portion of decomposition (explained= 0.007 and unexplained=0.182). Therefore, most part of the wage difference is not due to the difference in productivity. The major portion is from the intercept of the unexplained component which may correspond to unobserved parameters or discrimination.

Using twofold decomposition results, let assume that the unexplained component of the wage-gap happens due to male-female wage discrimination and the coefficient of pooled regression is considered non-discriminatory. Therefore, coefficients of unexplained component indicate that 42% of unexplained wage-gap originated from discrimination in favor of male workers ($\hat{\beta}_M - \hat{\beta}_R = 0.076$) and 58% of it is due to discrimination against female workers ($\hat{\beta}_R - \hat{\beta}_F = 0.106$).

The unexplained part of decomposition can be further decomposed. Let $\beta_M = \beta^* + \delta_M$ and $\beta_F = \beta^* + \delta_F$ with δ_M and δ_F as group-specific discrimination parameter vectors (positive or negative discrimination, depending on the sign). U can be expressed as

$$U = E(X_M)'\delta_M - E(X_F)'\delta_F \tag{4}$$

The unexplained component of differentials can be subdivided into a part, $U = E(X_M)'\delta_M$ that measures discrimination in favor of male workers and $U = -E(X_F)'\delta_F$ that measures discrimination against female workers (Jann 2008).

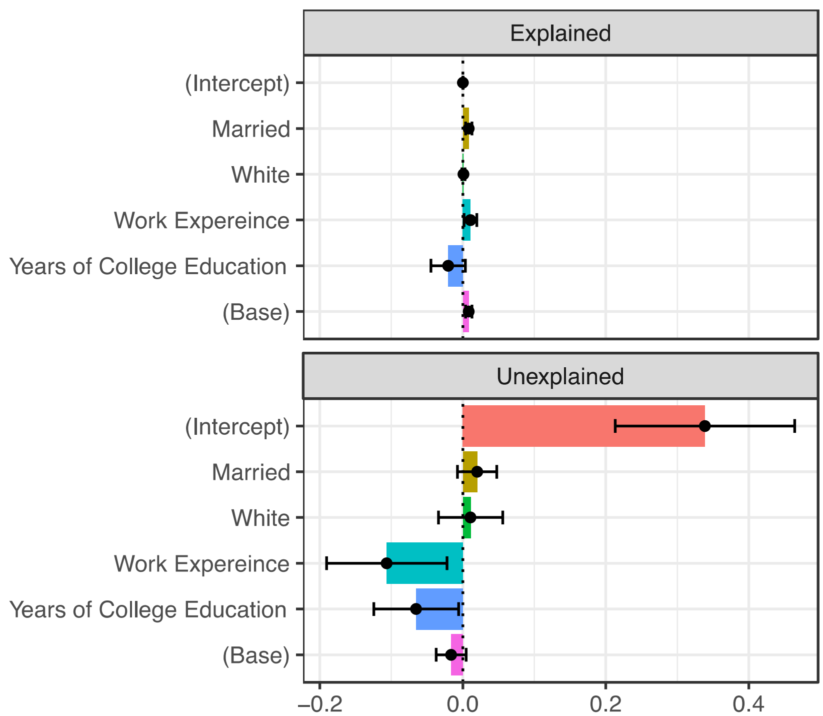

Figure 2. The explained and unexplained components of a twofold Blinder-Oaxaca decomposition of the female vs. male wage gap

Figure 3 visualizes the portion of the wage-gap that can be attributed to discrimination in favor of males, and how much is due to discrimination against the females. This figure shows the details on the individual contribution of predictors. Being married has a positive influence on the log-wage however, married women compared to married males are more discriminated. The discrimination for women with no college education and new to work tends to be even worse. However, an immense portion of the unexplained wage-gap lies within an intercept indicating there are unobserved variables that may explain the wage-gap.

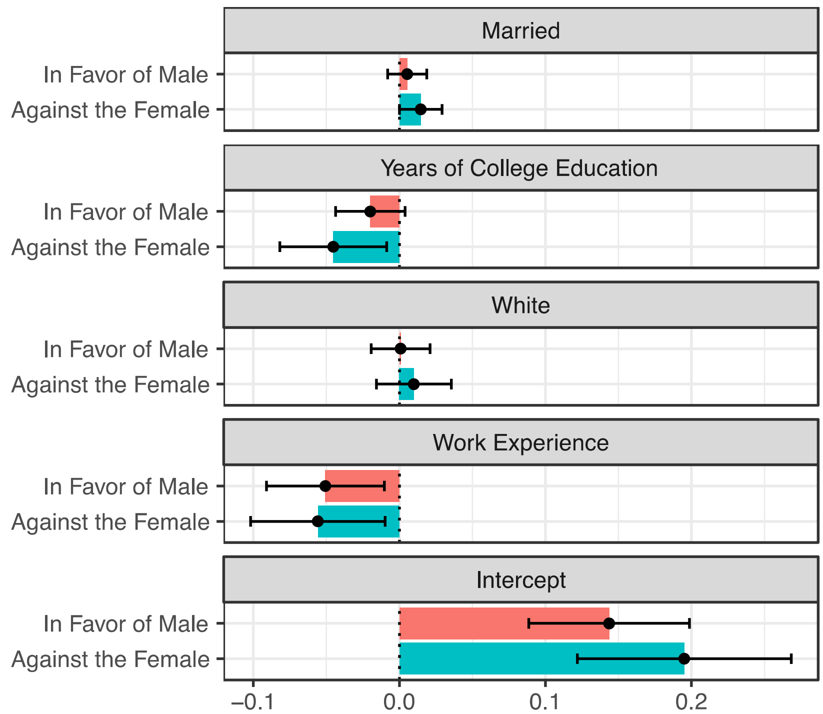

Figure 3. The unexplained portion's discrimination sub-components in a twofold Blinder-Oaxaca decomposition of the female vs. male wage gap.

According to the above findings, more than 95% of male-female wage differences cannot be explained with human capital characteristics. In average female workers earn ($29.05/hour) which is 82.7% of average male workers earning ($35.13/hour). This is not much different from the United States average ratio (81.8% female: male wage ratio) [BLS report 2017]. This significant gap shows despite policies in place such as the California Equal Pay Act, the wage-gap has not changed significantly in the past two decades [Blau and Kahn 2017]. The results of this paper indicate that using human capital parameters to measure individuals productivity leaves the major portion of the gap unexplained which might be due to unobserved predictors derived from culture, society or discrimination (Oaxaca 1973). Researchers have studied other parameters such as valuing money/work parameters or women being less risk avert (Fortin

2008). However, these parameters do not result in different productivity. Therefore we can conclude that the difference in the wage of men and women in the San Francisco Bay Area has rooted in characteristics different than workers productivity and job performance.

References

Becker, Gary. "S. Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education." (1964) New York, Colombia University Press.

Blau, Francine D., and Lawrence M. Kahn. "The gender wage gap: Extent, trends, and explanations." Journal of Economic Literature 55.3 (2017): 789-865.

Blinder, Alan S. "Wage discrimination: reduced form and structural estimates." Journal of Human resources (1973): 436-455.

Fortin, Nicole M. "The gender wage gap among young adults in the united states the importance of money versus people." Journal of Human Resources 43.4 (2008): 884-918.

Hlavac, Marek. "oaxaca: Blinder-Oaxaca decomposition in R." Available at SSRN 2528391 (2014).

Jann, Ben. "The Blinder–Oaxaca decomposition for linear regression models." The Stata Journal 8.4 (2008): 453-479.

Kunze, Astrid. "Gender wage gap studies: consistency and decomposition." Empirical Economics 35.1 (2008): 63-76.

Kunze, Astrid. "The evolution of the gender wage gap." Labour Economics 12.1 (2005): 73-97.

Neumark, David. "Employers' discriminatory behavior and the estimation of wage discrimination." Journal of Human resources (1988): 279-295.

Oaxaca, Ronald. "Male-female wage differentials in urban labor markets." International economic review (1973): 693-709.

US Department of Labor, Bureau of Labor Statistic “Highlights of women's earnings in 2017.” available at report 1075 (2017)