Title: Do Women Still Earn Less than Men after College Graduation:

Evidence from the Baccalaureate and Beyond Longitudinal Study 1993 Cohort

1. Context

Even though women have continuously caught up with men in education attainment and labor market participation since the 1970s (Goldin et al., 2006), the wage gap between men and women still universally exists today (Blau and Kahn, 2017). Do female college graduates still earn less than their male counterparts, if men's and women's "profiles" of observed productivity-related characteristics are statistically adjusted to be equivalent? To answer this research question and better understand the current gender wage gap, I introduce a novel propensity score stratification method for gender wage gap decomposition. This new method overcomes certain limitations of the traditional Blinder-Oaxaca decomposition method, and provides an example of validly applying propensity score-based methods (mostly used in causal settings) to gender wage gap decomposition, a non-causal setting.

2. Objective

The objective of this study is to explore whether gender wage gap still exist after the joint distribution of the observed productivity-related characteristics of men and women are statistically adjusted to be equivalent? In other words, should women had equivalent observed productivity-related characteristics as men, will they earn the same level of wage as men?

3. Setting

The data come from Baccalaureate and Beyond Longitudinal Study (B&B) that are conducted by the National Center for Education Statistics in the Institute of Education Sciences in United States between 1993 and 2003.

4. Population

The sample population is the 11,192 college graduates followed by the B&B study. The subjects for this longitudinal study was drawn from the National Postsecondary Student Aid Study to be a nationally representative sample of the 1993 cohort of college graduates from Title IV postsecondary institutions in United States.

4. Data Collection

The B&B study followed 11,192 college graduates for ten years and gathered information on these students' demographics, education experiences, employment trajectory, family formation, and other aspects of their lives through multiple waves. Specifically, data are collected in the senior year of these college students, the first, the fourth and the tenth year after their graduation.

5. Data Analysis

I use a novel method that augments the Blinder-Oaxaca framework (Blinder, 1973; Oaxaca, 1973) and applies propensity score stratification method to gender wage gap decomposition. Propensity score based methods are advantageous in reducing the multi-dimensional complexity of the covariates and are mostly used in causal setting. However, although gender is not the "treatment" as in the causal setting, the relationships between gender, covariates and wages can be interpreted in a way analogous to treatment, covariates and outcomes in the causal setting. Note that there is no causal claim, as the wage gap decomposition here is still a descriptive comparison of wage between men and women.

Essentially, the propensity score $e(X) = \Pr(N = f|X)$ is a balancing score. It is a function of the observed covariates X and has the property of "balancing" the joint distribution of these covariates between the two mutually exclusive groups, denoted by N:

$$X \perp N \mid e(X).$$

Given propensity score $e(X)$, the joint distributions of covariates X between the binary groups N are equated (Rosenbaum and Rubin, 1983). Under the BO framework, the "explained" component of the gender wage gap are the part that equal productivity and labor supply will lead to equal pay, where the productivity-related characteristics predicts the "explained" wage. In other words, the part of the gender wage differential that can be removed by making the observed productivity-related characteristics equivalent between men and women is the "explained" component of the wage gap. Any remaining part of the gender wage differential is the "unexplained" component of the wage gap.

As the decomposition analysis gender wage gap is non-causal, e(X) is hard to conceptualize. I propose to interpret it as the probability of being a woman given certain "profile" of the observed productivity-related characteristics. Suppose the employers in the market can observe all productivity-related characteristics X of employees and has the information on the proportion of women given the observed "profile" of these characteristics, which is e(X). However, the employers cannot observe the gender of a certain employee and can only guess based on the information. A labor market without gender discrimination will ignore the information on the conditional proportion of gender and will determine the wage levels purely according to the observed "profile" of productivity-related characteristics. However, a labor market with some degree of gender discrimination will consider the conditional proportion of gender and will assign different wages based on the conditional probability of this "profile" being a certain gender. The above rationale enables the application of propensity score based methods to the gender-based decomposition analysis without causal claims.

6. Results

The analytic results (Table 1) show that after adjusting all the covariates through propensity score stratification, being a woman means earning 12.6% less of the men's annual salary. In other words, after adjusting for the all the observed human capital characteristics and work-related experiences, the women-to-men wage ratio is 87.4% and statistically significant different than 100%, which is larger than the unadjusted women-to-men wage ratio (70.78%).

The results of the within-stratum gender difference in 2003 annual salary for all strata (Table 2 and Figure 1) show that the gender difference universally exists in all strata, and the majority of the gender ratios are statistically significant different than 100%. Across strata, average annual salary is generally decreasing from stratum 1 to stratum 21, that is, if the "profile" of observed covariates is more likely to be men's, the average annual salary is higher.

7. Conclusions

In conclusion, the observed gender difference in pathways into labor market cannot explain all the gender wage gap among the 1993 cohort of college graduates. However, due to the limitation of data, gender difference in psychological traits and job preferences are unobserved. Hence, the results can be viewed as the "upper bound" of the discriminatory gender wage gap before further analysis accounts for more explanatory variables. Results of gender wage gap of special populations, such as racial minority groups, first-generation students, low-income students are forthcoming.

Tables and Figure:

Table 1. Regression Results

VARIABLES	ln(wage03)
Gender (Female=1)	-0.126***
	(0.0227)
Constant	10.78***
	(0.0198)
Observations	8,543
R-squared	0.008

Table 2. Within-Stratum Difference in 2003 Annual Salary between Women and Men

	Wage Ratio	Average Wage	Women			Men		
Stratum			N	Mean	SD	N	Mean	SD
1	92%	60994	38	56599	0.08	484	61348	0.03
2	85%**	64174	71	55741	0.06	450	65634	0.02
3	76%***	60909	88	48784	0.06	433	63829	0.03
4	89%**	58134	140	53197	0.05	381	60084	0.03
5	84%***	53491	170	47692	0.05	352	56804	0.04
6	88%**	53217	192	49004	0.04	329	55770	0.04
7	85%**	53943	203	48995	0.05	318	57723	0.05
8	84%***	50915	239	46306	0.04	282	55380	0.04
9	92%	50458	280	48400	0.05	242	52746	0.07
10	83%**	45868	301	42457	0.05	220	51415	0.07
11	81%***	45154	307	41177	0.04	214	50881	0.06
12	93%	44258	317	42979	0.05	204	46207	0.05
13	83%**	43057	3337	40351	0.03	185	48463	0.05
14	92%	42064	373	41044	0.03	148	44582	0.07
15	90%*	38479	402	37503	0.03	119	41900	0.06
16	85%**	39967	413	38690	0.03	108	45413	0.06
17	89%	36565	430	35829	0.04	92	40147	0.06
18	117%	35953	458	36645	0.03	63	31300	0.17
19	110%	33125	466	33504	0.04	55	30325	0.19
20	88%	32595	482	32283	0.03	39	36829	0.10
21	68%***	30772	508	30493	0.04	13	45168	0.12

Note: SD display the standard deviations of log(wage03), the logarithms of 2003 annual salary of men or women within strata. *** p<0.01, ** p<0.05, * p<0.1

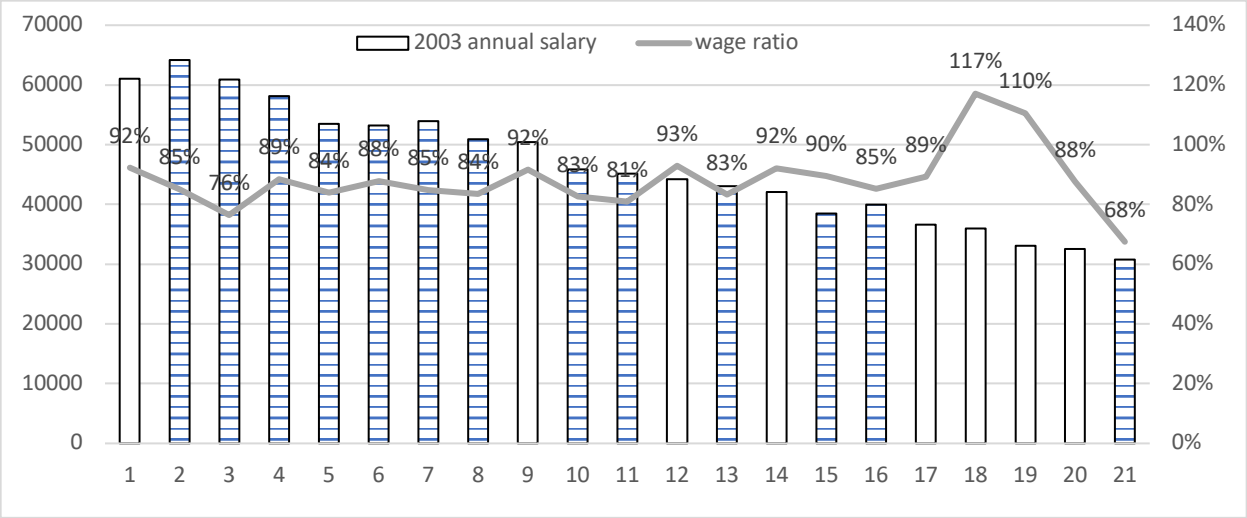

Figure 1. 2003 Annual Salary and Wage Ratio across Strata
(Pattern filled bars mean gender wage gap is statistically significant.)

Reference:

Blau, F. D., & Kahn, L. M. (2017). The gender wage 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*, 436-455.

Goldin, C., Katz, L. F., & Kuziemko, I. (2006). The homecoming of American college women: The reversal of the college gender gap. *Journal of Economic Perspectives*, *20*(4), 133–156.

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

ROSENBAUM, P. R., & RUBIN, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41–55.