

Gender Wage Gap in the US

In this segment, we will consider the case study of the gender wage gap in the United States. Here we will ask and answer the following question:

What is the difference in predicted wages between men and women with the same job-relevant characteristics?

In the previous case study (Predicting Wages), we worked with the US current population survey data for the year 2012.

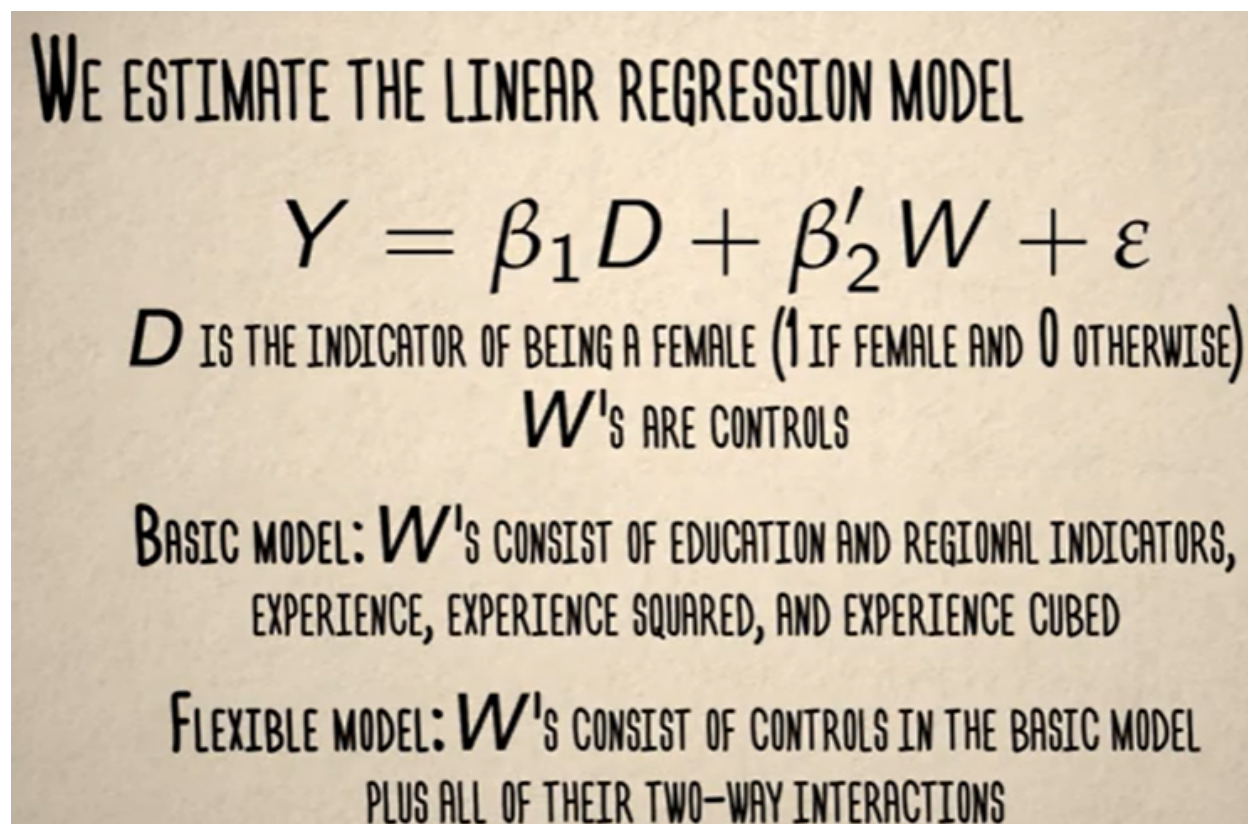
We first take a look at the following descriptive statistics for the sub-samples of single men and single women with educational attainment (highest level of education) equal to high school or some college. The mean hourly wage, as we see in the table below, is around \$16 for men and nearly \$15 for women. So there is a difference of around \$1, without controlling for job-relevant characteristics.

	Male means	Female means
Wage	16.12	14.72
Experience	13.58	13.04
College graduate	0.35	0.41
Some college	0.30	0.35
High school graduate	0.34	0.24
Midwest	0.28	0.29
South	0.24	0.26
West	0.22	0.20
Northeast	0.26	0.26

If we take a look at some of these characteristics, we see that **on average, the male workers have more experience, but the female workers are more likely to have college degrees or some college education.** The geographical distribution for men and women is, however, similar.

Now, we proceed to answer our question.

We estimate the linear regression model that we have introduced. Our outcome Y is the hourly wage and our target regressor D is the indicator of being female. Our controls W 's, are the job-relevant statistics.



WE ESTIMATE THE LINEAR REGRESSION MODEL

$$Y = \beta_1 D + \beta_2' W + \varepsilon$$

D IS THE INDICATOR OF BEING A FEMALE (1 IF FEMALE AND 0 OTHERWISE)

W 'S ARE CONTROLS

BASIC MODEL: W 'S CONSIST OF EDUCATION AND REGIONAL INDICATORS, EXPERIENCE, EXPERIENCE SQUARED, AND EXPERIENCE CUBED

FLEXIBLE MODEL: W 'S CONSIST OF CONTROLS IN THE BASIC MODEL PLUS ALL OF THEIR TWO-WAY INTERACTIONS

We consider two specifications for the controls:

1. **In the basic model**, W 's consist of education and regional indicators, as well as other regressors like experience, experienced squared, and experienced cubed.
2. **In the flexible model**, W 's consist of controls in the basic model, in addition to all of their two-way interactions.

We will now go to the **empirical results**.

In the following table, we see the estimated regression coefficient, its standard error and the 95% confidence interval for both the basic and flexible regression models.

	Estimate	Std Error	Confidence Interval
	$\hat{\beta}_1$	$\sqrt{\hat{V}/n}$	95%
Basic reg	-1.83	0.42	[-2.66 -0.99]
Flex reg	-1.88	0.42	[-2.71 -1.05]

The results for the two regression models are very close to each other.

THE ESTIMATED GENDER GAP IN HOURLY WAGE IS ABOUT $-2\$$ WITH A
CONFIDENCE INTERVAL THAT RANGES FROM ABOUT $-2.7\$$ TO $-1\$$

THIS MEANS THAT WOMEN GET PAID $2\$$ LESS PER HOUR ON AVERAGE
THAN MEN, CONTROLLING FOR EXPERIENCE, EDUCATION, AND GEOGRAPHICAL REGION

We conclude that the difference in the hourly wage for men and women who have the same recorded characteristics is **both statistically and economically significant**.

To sum it up, we apply the ideas we discussed before and learn about the gender wage gap. The gender wage gap may partly reflect genuine discrimination against women in the labor market.

It may also partly reflect the so-called **Selection Effect**, namely that women are relatively more likely to take on occupations that pay somewhat less. School Teaching is one such example.