Regressions with Interaction Effects

	Dependent variable:			
	wages			
	(1)	(2)	(3)	(4)
Constant	22,969***	9,584***	6,323***	7,847***
	(493)	(365)	(339)	(421)
female	6,711***		6,598***	3,565***
	(698)		(309)	(595)
years:female				200***
				(34)
years		1,104***	1,102***	1,001***
		(21)	(17)	(24)
Observations	1,000	1,000	1,000	1,000
Note:				*p<0.1;* ***p<0.05;*** p<0.01

Model 1: Do men and women earn different wages, on average?

Model: $Wages = b_0 + b_1 * Female$

Test: If b0 is significant, Men's wages are different than zero. If b1 is significant, Women's wages are different than Men's.

Note: This is an unconditional average, so it might be explained by other factors like differences in experience between men and women.

Model 2: What is the wage gain related to an extra year of experience?

Model: $Wages = b_0 + b_1 * Years$

Test: If b1 is significant then it is different than zero, experience does impact wages.

Model 3: Do men and women have different initial wages at the start of their careers?

Model: $Wages = b_0 + b_1 * Years + b_2 * Female$

Test: If b2 is significant then the Female intercept (b0+b2) is different than the Male intercept (b0).

Model 4: Are the gains in wages related to experience the same for men and women?

Model: $Wages = b_0 + b_1 * Years + b_2 * Female + b_3 * Years * Female$

Test: If b3 is significant then the slope for Women (b1+b3) is different than the slope for Men (b1).

Note: If b3 is not significant it is better to use the model with one slope for both groups.

Use Model 4 to answer the following questions:

(1) **How much do men earn in their first job (zero years of experience)?**

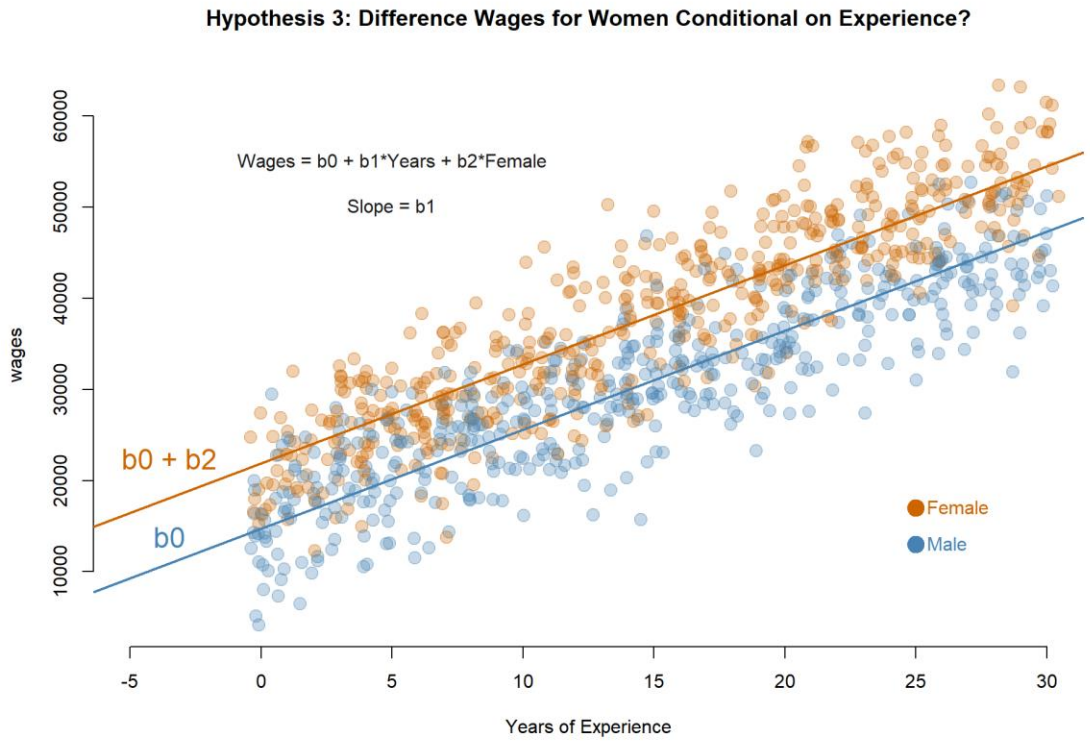
Male is the omitted category, so it is represented by the intercept (constant b0) and setting years equal to zero we get:

Wages = 7,847 + (0)(1,001) = **$7,847**

(2) **How much do women earn in their first job?**

To find the female starting wage, since female is the dummy variable included in the model we would use the coefficient to adjust the baseline intercept:

Wages = 7,847 + 3,565 + (0)(1,001) = **$11,412**



(3) **What is the average raise men receive each year?**

Raises are represented by the increase in salaries by year, so the slope associated with years, or the average raise, would be:

$1,001 per year

(4) **Do men and women receive different raises? How do we know?**

When we include the interaction of the dummy variable and the years, this is a test of whether the slope associated with years is different for the groups (men and women in this case). The answer will never be mathematically zero, so we need to know whether we can interpret the difference as meaningful, which is what we learn from the statistical significance of the **years:female** interaction.

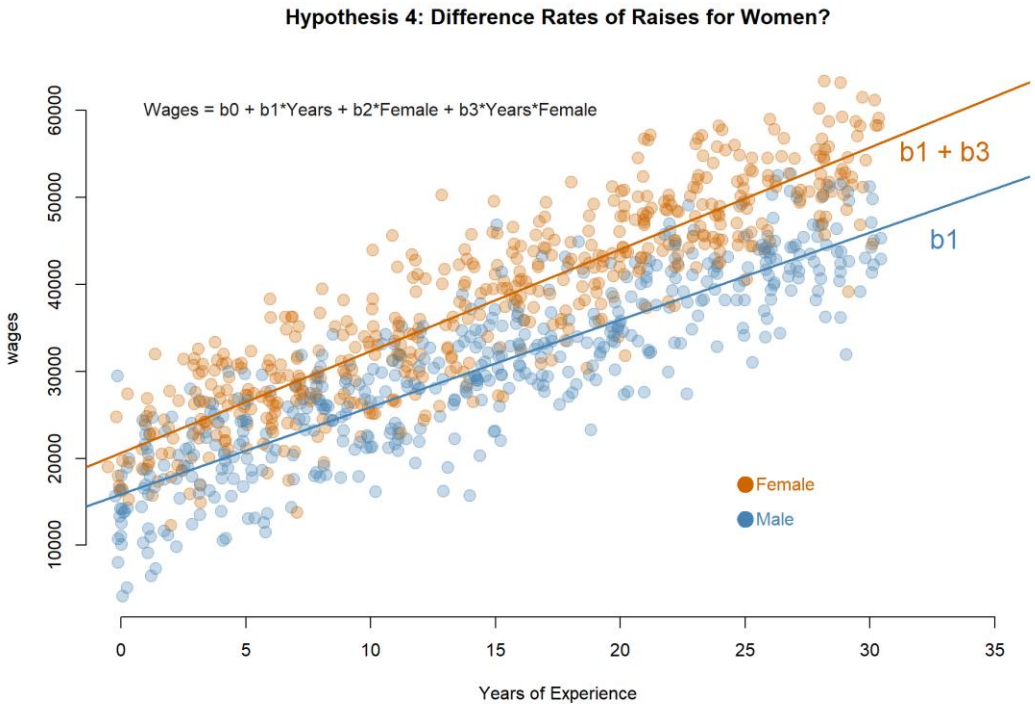
Since the slope has three stars by it, that signifies a high significance level. We conclude that average wage increases are different between men and women.

Alternatively, if we build the confidence interval we see: 200 +/- 1.96(34) would not contain zero.

(5) **What is the average raise women receive each year?**

The wage increase for women will then be years + years:female, or:

1,001 + 200 = **$1,201 per year**



Note that it is always a good idea to write down the full equation and then figure out which dummy variables would be 0 or 1 for each question, then do the math.

For example, for the wage offer for the first job for men, we would write:

Wages = 7,847 + (3,565)(female) + (1,001)(years) + (200)(years:female)

Since men are the omitted group in this model, we replace each female dummy with zero (1=female, 0=male). And since the question is asking for starting wages, we replace years of experience with zero:

Wages = 7,847 + ~~(3,565)(0)~~ + ~~(1,001)(0)~~ + ~~(200)(0)~~

And thus we drop everything except the intercept.