

# The Impact of Educational Attainment and Gender on Earnings in Puerto Rico

By: Joe Thomson, Majorie Blanco, Haodi Tu  
In partial fulfillment of the requirements for Econ 5300  
Professor Bridget Heidemann  
Winter Quarter 2019

## Introduction

The United States regularly implements policies and programs to promote the economic welfare of its citizens. Economic inequality can lead to social and political unrest so it is in the interest of the country to identify whom policies should be targeted towards and what form those policies should take. This paper examines two related questions using a sample of residents from Puerto Rico. First, how do earnings vary by education, and second, how does the premium for education vary by gender? Answering these two questions together provides information about how policies or programs related to education can be tailored for maximum effect in light of a person's gender. We use an ordinary least squares (OLS) model stratified by gender to analyze our data. What we find is that higher levels of educational attainment are positively correlated with earnings and that the premium for men is higher than for women with the exception of those who obtain professional degrees.

## Econometric Model and Estimation Method

The dependent variable in our model is the natural logarithm of annual earnings. Earnings are defined as all wages, salaries, commissions, bonuses, tips, and self-employment income during the past 12 months. Log-level models are appropriate when percentage differences in the dependent variable (earnings) are more relevant than unit differences. It is also helpful in correcting skewness in regression models. It makes sense from a theoretical perspective to look at percentage differences in earnings rather than unit differences when one considers the large differences in earnings between careers that require the same level of education. For instance, university professors, while often holding doctoral degrees, would be expected to make significantly less than physicians. Hence, in unit terms, we would expect to see all earnings levels represented at all education levels. The logarithm of earnings gives us a more meaningful comparison of the relative effect of education.

Dummy variables for educational attainment: no college (the reference category), some college, associate's degree, bachelor's degree, master's degree, professional degree, and doctoral degree.

Dummy variables for race: white (the reference category), black, other. The overwhelming majority of Puerto Ricans are white, with black being the only sizable minority. Other racial/ethnic groups were judged too small to evaluate on their own.

Dummy variables for marital status: married (the reference category), widowed, divorced, separated, never married.

The model also includes variables for age and age squared. It is appropriate to include age squared as its quadratic form fits with the logarithm of earnings.

If Puerto Rico was larger geographically, it might make sense to control for urban versus rural populations. Puerto Rico is also in a unique position as large numbers of youth leave the territory to pursue opportunities on the U.S. mainland.<sup>1</sup> The impact of the out-migration of educated youth on the Puerto Rican economy and its interaction with educational attainment and gender could be an area for further research.

## Data

Our data comes from the 2016 American Community Survey for Puerto Rico. The full data set includes approximately 30,000 observations as a representative sample for all of Puerto Rico. It collects data on earnings, employment, and various personal and demographic characteristics such as education, marital status, race, and age, all of which factor into our analysis. We limit our sample to full-time year around working age individuals, defined as those who worked more than 35 hours per week for at least 40 weeks during the year, and are between the ages of 18 and 64. We also set an earnings range of \$10,000 to \$159,000. The upper limit is the top coded amount in the survey, and the lower limit serves to remove outlying earnings that intuitively don't make sense for a full-time worker. These criteria reduce our final sample to 5,194 with a nearly even split between genders. These selection criteria introduce some limitations to our analysis. For instance, there may be a relationship between the ability or desire to find full time work and

---

<sup>1</sup> <http://www.pewresearch.org/fact-tank/2015/10/14/puerto-ricans-leave-in-record-numbers-for-mainland-u-s/>

education or gender which our analysis does not address. However, given the many reasons individuals decide not to work full-time (such as doing valuable but non-revenue generating labor like caring for family members), we felt it was appropriate to limit our sample population in this way.

Tables 1 and 2 show descriptive statistics for total earnings by gender. As expected, mean earnings for men are higher than for women by approximately \$1,500, but they also have a larger standard deviation. Maximum earnings for men are \$125,000 while maximum earnings for women are \$98,000. Tables 3 and 4 show earnings by gender and education. Women's average earnings lag man at each level, with the exception of those who obtain professional degrees where the means between genders are equal, but median earnings for men are still greater than those for women. Tables 5 and 6 look at earnings by gender and marital status. Interestingly while married men earn the most among men, widowed women earn the most among women. Overall, our descriptive statistics confirm our expectations that higher levels of education lead to higher earnings, and that men earn more than women on average.

	<b>Table 1: Descriptive statistics</b>				
	<b>SEX: Male (N = 2,627)</b>				
	<b>Mean</b>	<b>Median</b>	<b>S.D.</b>	<b>Minimum</b>	<b>Maximum</b>
<b>Total person's earnings</b>	30049.84	24000.00	18092.58	10000.00	125000.00
<b>Log Total person's earnings</b>	10.15800	10.15800	0.5381167	9.21034	11.73607
<b>Age</b>	42.8	43	11.1	18	64

	<b>Table 2: Descriptive statistics</b>				
	<b>SEX: Female (N = 2,567)</b>				
	<b>Mean</b>	<b>Median</b>	<b>S.D.</b>	<b>Minimum</b>	<b>Maximum</b>
<b>Total person's earnings</b>	28489.81	24000.00	16127.88	10000.00	98000.00
<b>Log Total person's earnings</b>	10.12812	10.08581	0.4910962	9.21034	11.49272

Age	43.5	44	10.5	19	64
-----	------	----	------	----	----

Table 3: Descriptive statistics Total person's earnings by Education					
	SEX: Male (N = 2,627)				
	Mean	Median	S.D.	Minimum	Maximum
High school diploma	22807.70	18600.00	12562.66	10000.00	125000.00
Some college	26160.31	22000.00	14147.83	10000.00	93000.00
Associate's degree	26892.20	22000.00	14708.86	10000.00	90000.00
Bachelor's degree	36505.19	32000.00	19182.20	10000.00	98000.00
Master's degree	42688.89	39500.00	21542.63	10000.00	98000.00
Professional degree	49276.92	48800.00	22230.10	13000.00	100000.00
Doctorate degree	65140.48	66000.00	20699.35	24000.00	96000.00

Table 4: Descriptive statistics Total person's earnings by Education					
	SEX: Female (N = 2,567)				
	Mean	Median	S.D.	Minimum	Maximum
High school diploma	20201.18	17000.00	9551.18	10000.00	80000.00
Some college	22676.00	18850.00	12125.99	10000.00	85000.00
Associate's degree	23163.35	19950.00	11326.66	10000.00	87000.00
Bachelor's degree	30312.91	27000 .00	15738.76	10000.00	98000.00
Master's degree	35810.46	32100 .00	17399.43	10000 .00	95000.00
Professional degree	49036.07	44000 .00	21644.73	10400.00	91000.00
Doctorate degree	52800.00	50000 .00	22202.05	17900.00	90000.00

Table 5: Descriptive statistics Total person's earnings by Marital Status					
	SEX: Male (N = 2,627)				
	Mean	Median	S.D.	Minimum	Maximum
Married	33085.87	27000.00	19133.67	10000.00	125000.00
Widowed	23778.57	27750.00	11090.69	10000.00	40000.00

<b>Divorced</b>	30628.89	24000.00	18736.58	10000.00	100000.00
<b>Separated</b>	28359.80	26400.00	14330.31	10000.00	65000.00
<b>Never married</b>	24491.74	19000.00	14470.40	10000.00	90000.00

<b>Table 6: Descriptive statistics Total person's earnings by Marital Status</b>					
	<b>SEX: Female (N = 2,567)</b>				
	<b>Mean</b>	<b>Median</b>	<b>S.D.</b>	<b>Minimum</b>	<b>Maximum</b>
<b>Married</b>	29725.53	25000.00	16837.41	10000.00	95000.00
<b>Widowed</b>	33192.16	27600.00	18679.94	10600.00	98000.00
<b>Divorced</b>	28957.19	25000.00	15898.88	10000.00	95000.00
<b>Separated</b>	26678.57	24150.00	12984.70	10000.00	78000.00
<b>Never married</b>	25930.95	21600.00	14795.41	10000.00	94000.00

## Results

Table 7 shows the estimates of our log earnings equations for males and females. The formula takes the following form:

$$\text{Log (Earning| male/female)} = \beta_0 + \text{Widowed} * \beta_1 + \text{Divorced} * \beta_2 + \text{Separated} * \beta_3 + \text{NeverMarried} * \beta_4 + \text{RaceBlack} * \beta_5 + \text{RaceOther} * \beta_6 + \text{SomeCollege} * \beta_7 + \text{Associate} * \beta_8 + \text{Bachelor} * \beta_9 + \text{Master} * \beta_{10} + \text{Professional} * \beta_{11} + \text{Doctoral} * \beta_{12} + \text{Age} * \beta_{13} + \text{Age}^2 * \beta_{14}.$$

The results indicate that earnings are positively associated with educational attainment for men and women in Puerto Rico, and that men earn more than women on average. Approximately 25% of female earnings are explained by our regression model, and the model has a high F-value (F-stat = 60.4), indicating the model as a whole is statistically significant. 28% of male earnings are explained by our model, and it is also statistically significant (F-stat = 72.1). Both models have low residual standard errors (female: 0.427 on 2552 degree of freedom; male: 0.458 on 2612 degree of freedom), meaning that they fit subsample data well and suggesting strong evidence that our models have predictive ability.

In our combined model, with a subsample of female, controlling for marital status, race, education and age, females earn on average about 13.09% less than males at a statistically significant level ( $p < 0.01$ ).

The estimated earnings premium for each level of educational attainment relative to the lowest level— no college degree—is higher by each increasing level of education attainment among men and women. For example, women who have some college education earn on average 13.5% more than women who have no college education after controlling for marital status, race and age. Women who have doctoral degrees earn on average about 149% more than women who have no college education after controlling for other variables. Both coefficients are statistically significant ( $P < 0.05$ ). Similar to women, men’s earnings increased with higher educational attainment with one exception. Men who have some college education earn about 15.7% more than men who have no college education, but men who have an associate’s degree about 14.6% more than men who have no college education, at statistically significant level ( $p < 0.05$ ).

As shown in Table 7, the estimated earnings premium for the same level of educational attainment is higher for men than for women in most cases and is statistically significant at the 1% level in all cases. The exception is that women with professional degrees have a higher premium over those with no college education than men.

In our subsample of male, predicted earnings increase at a decreasing rate within the observed age range of 18 to 64 years. The average age for maximum earnings is 53. The coefficients on age and its square are statistically significant.

Table 7: Estimated Log (Earnings) Equations			
<i>Dependent variable:</i>			
Log (Earnings)	Combined (1)	Female (2)	Male (3)
<b>Female</b>	-0.140*** (0.013)		
<b>Widowed</b>	0.039 (0.061)	0.126* (0.066)	-0.244* (0.130)

<b>Divorced</b>	-0.042** (0.018)	-0.008 (0.023)	-0.078*** (0.028)
<b>Separated</b>	-0.047 (0.043)	-0.038 (0.056)	-0.048 (0.067)
<b>NeverMarried</b>	-0.087*** (0.016)	-0.031 (0.021)	-0.145*** (0.023)
<b>RaceBlack</b>	-0.031* (0.017)	-0.047** (0.023)	-0.018 (0.025)
<b>RaceOther</b>	-0.060*** (0.016)	-0.085*** (0.022)	-0.037 (0.024)
<b>SomeCollege</b>	0.148*** (0.020)	0.127*** (0.031)	0.156*** (0.026)
<b>Associate</b>	0.150*** (0.019)	0.139*** (0.029)	0.146*** (0.026)
<b>Bachelor</b>	0.416*** (0.017)	0.390*** (0.025)	0.441*** (0.024)
<b>Master</b>	0.565*** (0.026)	0.559*** (0.033)	0.562*** (0.044)
<b>Professional</b>	0.816*** (0.051)	0.870*** (0.068)	0.734*** (0.074)
<b>Doctoral</b>	0.962*** (0.045)	0.912*** (0.066)	1.016*** (0.060)
<b>AGEP</b>	0.032*** (0.004)	0.016*** (0.006)	0.045*** (0.006)
<b>I(AGEP * AGEP)</b>	-0.0003*** (0.0001)	-0.0001 (0.0001)	-0.0004*** (0.0001)
<b>Constant</b>	9.115*** (0.091)	9.291*** (0.130)	8.890*** (0.128)
<b>Observations</b>	5,194	2,567	2,627
<b>R<sup>2</sup></b>	0.259	0.249	0.279
<b>Adjusted R<sup>2</sup></b>	0.257	0.245	0.275



<b>Residual Std. Error</b>	0.444 (df = 5178)	0.427 (df = 2552)	0.458 (df = 2612)
<b>F Statistic</b>	120.803*** (df = 15; 5178)	60.487*** (df = 14; 2552)	72.132*** (df = 14; 2612)
<b>Note:</b>	Estimates are obtained using generalized estimating equations. Robust standard errors are in parentheses. *p<0.10 **p<0.05 ***p<0.01		

### Conclusions

If Puerto Rico were interested in closing the earnings gap between men and women, it would be well advised to encourage women to pursue higher levels of education in general. The fact that the premium for women's education lags that for men, however, suggests there may be other factors that either cause women to pursue less lucrative careers or be paid less in the same careers as men. This could be due to any number of social, cultural, or religious factors that would need to be addressed in those venues. It could also be the result of gender prejudice. That being said, our results indicate that women benefit more than men from receiving professional degrees. One explanation of this higher premium could be that the standardized competencies conferred by these degrees (such as law or medicine) counteract gender prejudice. Another explanation is that these happen to correspond to several of the more lucrative careers women in Puerto Rico might pursue. Either way, a focused effort to move women through the professional degree stage could be an effective way to begin closing the gender earnings gap, and further study of this possibility is recommended.

### Contributions

All members were involved in meetings to discuss sample selection criteria, model specification, and the interpretation of results. Other contributions were as follows:

Marjorie Blanco: R coding

Haodi Tu: Analyzing of results and data formatting

Joe Thomson: Writing and editing of final paper and conclusions