

Meghan Peters, Tommy Chen  
Professor Shukla - Section I1  
Project B  
April 18, 2017

### Gender Salary Gap: The Role of Variables

It is an indisputable fact that women as a whole earn less on average than men in the American workforce. There are varied reports on exactly how much less women tend to earn, although it is generally stated as 83 cents to a man's dollar, and an array of variables that may contribute to the pay gap. It is unlikely that as a whole, all employers consciously and tenaciously chose to pay women lesser salaries than men; the purpose of this paper is to ascertain the possible sources of the gender wage gap within a specific industry. If the origins of the wage gap can be identified, employers and policy makers will be enabled to take proactive measures that address and continue to close the wage gap in an efficient manner. The literature of legal documents in public policy is incredibly articulate and specific. Policies must be meticulously laid out in order to ensure proper enforcement and clarification of the parameters of the law. By having specific information and regressions that address the various aspects of the gender wage gap, policy makers will be able to isolate the particular variables that have a strong correlation (implying a contribution) to the gender wage gap. Through this isolation, policy makers can write effective legislation that directly confronts the presumable sources of the gap. While most reports on the gender wage gap utilize a vast data base that encompasses a variety of individuals within multiple industries, by analyzing the gender wage gap in an isolated industry (real estate) a more precise conclusion is possible since all individuals are in a homogenous industry. Through the use of regressions, it is likely that the gender wage gap stems from a combination of factors, however there is no one specific cause.

The data utilized in this paper is derived from the 2013-2015 American Community Survey collected by the US Census Bureau. The sample is confined to workers who, at the least, have a Bachelor Degree in Business and work in a "professional occupation" as defined by the Census, in the real estate industry. The scope of the data is inclusive of 1,662 workers aged 25 to 60, who reported working 50- 52 weeks a year and have an income that surpasses the minimum wage. The dataset addresses the subsequent variables: age, usual hours worked weekly, annual salary, occupation, gender, state of residence, class of worker, degree subfield, advanced degree, marital status, and number of children in household. It is important to address the variables that are most strongly correlated to the largest apertures in the wages of men and women. The isolation of specific variables enables statisticians to make claims regarding the degree to which said variables contribute to the wage gap. Through the analysis of regressions and the results of Project A, it was determined that out of the dataset provided, the most influential variables to the gender wage gap are as follows: number of children in household, salary on gender, working hours, variable wage per hour, salary on age, advanced degree, and accounting as an occupation. Accounting was chosen as the occupation indicator due to its frequency in the data (43.56% of survey participants stated accounting as their occupation). Accountants have strict and specific requirements for entry into the field, ensuring an equal merit qualification basis. Additionally, accounting as an occupation had the most centralized degree subfield (66.58% of accountants listed accounting as their degree subfield); these factors lessen variation of the caliber of the individual and aid in production of an accurate analysis. The evidence produced in this paper has

been constructed from the analysis of regressions which have been calculated through excel functions. The regressions were run using information derived from the dataset collected by the US Census Bureau as previously stated. All references are made in relation to the defined dataset relevant to the real estate industry, and are based on the specific information yielded by the individuals who participated in the 2013-2015 American Community Survey. Although the data is specific to those individuals, the patterns that emerge are applicable, to a degree, to the entire industry and arguably to the gender wage gap as a whole. There must however be a disclaimer made that these are simply patterns used to analyze an underlying societal issue and is in no way indicative of an absolute truth or to be used as a definitive and conclusive evaluation of the wage gap.

To begin the process of understanding the sources of the gender wage gap, it is imperative to first establish a foundation for regressions and analysis. This is done in part by organizing the data into a table of descriptive statistics of earnings by gender and creating a histogram built on the information provided by the table. This information yielded an analysis that men earn more than women in every statistical category, other than a minimum in which men have recorded a lower salary than women. Male's salaries are also more varied than women's as they have a larger range and standard deviation. The median salaries of men and women are much closer in proximity than the mean salaries, suggesting men have more outliers supported by the fact they have a greater range. Men and women appear to earn similar amounts with similar frequency in the \$55,000-111,000 range. However, as salary increases, the frequency of women receiving higher salaries in regards to the frequency of men receiving higher salaries drops dramatically.

In order to achieve a more comprehensive analysis of the data and draw a more accurate conclusion, regressions are run using various variables to make assumptions regarding the relationships between these variables and their impact on the gender wage gap. This is done to avoid omitted variable bias which occurs when an important variable(s) is not accounted for and therefore forces the remaining variable(s) to be either over or undercompensated for, skewing the results to be inaccurate. Four regressions were run utilizing excel functions. The statistical process to achieve these regressions is initiated using the data analysis regression excel function. This requires the input of a Y-Range which defines the dependent variable and determines what you are attempting to predict. In the case of all four regressions, the Y-Range is "Annual Salary." Additionally, you must input an X-Range which is used to explain or predict the dependent variable. The X-Range is defined by the chosen independent variable(s) and are identified in the four regressions as stated below:

Regression 1: Gender (indicator variable)

Regression 2: Gender (indicator variable) and number of hours worked per week

Regression 3: Gender (indicator variable), number of hours worked per week, and accounting as an occupation (indicator variable)

Regression 4: Gender (as an indicator variable), number of children in a household, number of hours worked per week, age, the existence of an advance degree (indicator variable), and accounting as an occupation (indicator variable)

The four regressions evaluate the relationships between the independent variable(s) stated above and their influence on the dependent variable which, in every regression, is annual salary. The P-Values of each variable in Regressions 1-3 indicate high statistical significance and a pertinence

to the relevance of the regression. In regression 4 however, the P-Value of all variables, excluding only age and hours worked per week, are not indicative of statistical significance. This suggests they may not be apposite in consideration of the wage gap. This is particularly interesting because gender itself is no longer necessarily influential. In each regression, the coefficients were negatively correlated with the female gender, supporting the statement that women consistently earn less on average than men. The coefficients gradually increase as the regressions add more independent variables, indicating that the wage gap is lesser when genders are compared with intensified scrutiny. The increase of the regression coefficients from the first through to the fourth implies that the weight of gender on salary is less substantial when compared with supplemental factors. The  $R^2$  increases the regression's predictive power from being very dependent on unobserved factors in the first three regressions, to being determined strongly by observed factors in the fourth regression. This is a result of the incorporation of additional variables; where there are more observable factors, an increasingly comprehensive view of the wage gap is possible. A higher  $R^2$  increases the accuracy of the regression and insinuates the information is effective in encompassing the origins and severity of the wage gap.

It is impossible to make a conclusive statement about wage discrimination based on gender. Although the data used in this analysis is inclusive of a variety of individuals from diverse backgrounds, there are a plethora of possible variables as well as non-quantitative factors which are impossible to account for and may affect the wage gap. Due to the inaccessibility of perfect information, there is inevitable potential for alternative correct conclusions regarding the source of the wage gap. Analysis of the regressions do however strongly support the statement that women do earn less than men on average even with the Equal Pay Act in effect. The rift in salary between the genders decrease greatly as an increasing number of independent variables are factored in. This leads to an assumption that all else equal, the salary gap itself is not as substantial as one may think. This encourages the possibility that variance in salary may be attributed not necessarily to women being paid less, but instead that women as a whole are not provided equal opportunity to be in a position where, when comparing salary, all factors excluding gender are coequal. An additional variable that should have been considered in the dataset would be race/ethnicity. Aside from gender, race and ethnicity continue to be the highest contributors to discrimination, especially in the work place. If this had been accounted for, it is likely that there would be a much larger gap in the salaries of Caucasians and minorities rather than simply men and women, particularly Caucasian men and minority women.

Ethical Analysis:

It would be wise of the Board to address the gender wage gap utilizing a social justice framework that outlines ethics to be about fairness and justice. Through utilization of the veil of ignorance, the company can ensure their method of administering pay is fair by creating a system in which they determine salary, bonuses and other forms of financial compensation using a blind structure that only focuses on performance reviews and analysis of the contribution of the employee to the company. Additionally, the company should ensure their hiring methods provide equal opportunity to all genders. Various blind spots that may lead to poor decision making include motivated blindness. The reallocation of funds to fairly pay all employees means some individuals at the company will not be paid as much. It is easy to ignore the wage gap when you benefit. Conflict of interest applies the same mindset; to avoid this have someone unaffected in a designated role to evaluate pay. Other blind spots encompass ethical fading as gender pay gap is an issue that effects all of American society and economy based on historical cultural problems. The concern for fairness begins to fade because it is happening everywhere else, desensitizing the company to pay injustice. In group favoritism may also affect the company's ability to make good decisions. By diversifying the board to be inclusive of all genders, ethnicities, etc., a varied group will lessen the opportunity to choose pay methods that favor a specific group and discriminate against another