DePauw University
Econ 350C: Statistics for Economics and Management
Professor Ophelia Goma

A Study on the Causes of Gender Wage Gap

Tyler Schaefers and Bijay Ranabhat

Abstract:

In the United States, the gender wage gap is a pertinent measure of gender inequality. In this paper we will analyze the causes of gender wage gap in the United States, looking at variables measured in each individual state, including Washington D.C. This study consists of four independent variables: the percentage of women in poverty, percentage of women who have received a bachelor's degree, percentage of women in STEM, and the percentage of women in managerial or professional occupations. Through regression analysis, we will measure the effect of each of these variables on the gender wage gap percentage [(income of men-income of women)/income of men)*100], and the data we collected will highlight the pertinent cause(s) of this gap. Our data concludes that education plays a significant role in the gender wage gap. The percentage of women who have received a bachelor's degree is inversely correlated to the gender wage gap percentage. Additionally, our data shows that as the percentage of women in poverty shows an inverse relationship; however this variable does not show statistical significance. The percentage of women in STEM, and the percentage of women in managerial or professional occupations have a positive correlation to the gender wage gap percentage, but they are not statistically significant.

Introduction:

All over the world, gender inequality plays a major role in society's decision making process. In Mukesh Eswaran's book, *Why Gender Matters in Economics*, he states that most societal conventions have "been formulated by men . . . and women have no means of articulating their experiences because they have been denied the education and leisure needed for this" (Eswaran 61). This statement highlights patriarchal beliefs, displaying the roots of gender discrimination. What people fail to realize is that these beliefs still linger, even in first world countries. For our research we will be focusing specifically on the gender wage gap in the United States. The issue that we are addressing in our research is whether or not the gender wage gap is caused by gender discrimination or if there are other factors that contribute to the wage differences between men and women. The variables we will be focusing on in our research include: the percentage of women in poverty, percentage of women who have received a bachelor's degree, percentage of women in STEM, and the percentage of women in managerial or professional occupations in correlation to the wage gap percentage.

Research shows that on average, the income for men is significantly higher than women (see appendix B.5.). Poverty is one possible variable that effects the gender wage gap. According to May Cohen's research "As more and more women live without men, either by choice or necessity, women's poverty becomes more visible" (Cohen 950). The effects of poverty cause an unequal division of labor, "clustering of women in lower income jobs," placing many women in a difficult position (Cohen 951). Because of the distribution of labor, women tend to have a lower average income than men. The factor that combats unequal distribution of labor is the education attainment received by women. In Leslie McCall's study on "Gender and the New Equality", she explains the increasing wage gap between workers with a college education and those without (McCall). Although this is the case, the question of whether or not gender discrimination plays a major role in gender wage differences: “even if women and men have the same amount of education, the human capital of men tends to be higher, because they have, on average, more experience in the labor market.” (Eswaran 103). Because of this, our model takes into account the occupations that women hold. Job evaluation, traditionally a management tool, can be used to ascertain whether jobs of comparable value are paid equitably.” (Figart and Kahn 21). However there are still gender biases that take place in this research. In Figart and Kahn’s regression analysis, they measure the relationship between points and pay for men and women in similar paying jobs. The results show that the “female salary line is lower and flatter than the male line, indicating wage discrimination.” (Figart and Kahn 22-23).

With all of this research in mind, our goal is to figure out if the percentage of women in poverty, percentage of women who have received a bachelor’s degree, percentage of women in STEM, and the percentage of women in managerial or professional occupations are correlated to the wage gap percentage, or if we are living in a society where gender discrimination is still prevalent. Through regression analysis of these variables, we will compare our results to the research conducted by other economists, in order to get a better understanding of the causes of the gender wage gap.

The Model:

Our regression model-

$Y = B_0 + B_1(PRatex_1) + B_2(BDegreex_2) + B_3(SWorkersx_3) + B_4(Occupationx_4) + e$

Where:

Y = Gender wage gap percentage in each state

$PRatex_1$ = The poverty rates of women in each state

$BDegreex_2$ = Percentage of women with a bachelor's degree in each state

$SWorkersx_3$ = Women's share of STEM workers in each state

$Occupationx_4$ = Percentage of women employed in managerial or professional occupations in each state

By running a regression test on the model above, we will determine the correlation between the independent variables ($PRatex_1$, $BDegreex_2$, $Sworkersx_3$, $Occupationx_4$) and the dependent variable (Gender wage gap percentage in each state). The regression test on the model will explain the significance of the model as a whole. We expect to see a positive correlation between the poverty rates of women and the gender wage gap from our regression results: the lower the poverty rate is for women, the lower the gender wage gap percentage. Unlike the first independent variable, we anticipate that there will be an inverse correlation between the percentage of women with a bachelor's degree, women's share of STEM workers and percentage of women employed in professional occupations in relation to the gender wage gap percentage. This expectation should be correct if a rise in any of these three independent variables points toward the decline of the gender wage gap.

1) The first independent variable we used to describe our model is the percentage of women in poverty. According to May Cohen, poverty causes women to end up working at home or in low paying, low status jobs (Cohen 951). Based on this, we believe that states which yield a higher percentage rate of women in poverty will cause women to earn a lower income, thus the gender wage gap will be higher.
2) For the second independent variable, we looked at the percentage of women with a bachelor's degree. If the conventions are true, we believe that women with a bachelor's degree are likely to have similar wages to men with the same qualifications. Traditionally, women have been segregated into low paying jobs, but this convention has been rejected by women with college degrees as they compete with men for higher paying jobs (McCall, 235). This argument implies that it is more likely that the gender wage gap percentage will be lower in states where women have a high percentage of bachelor's degree attainment compared to the states where the percentage of women with a bachelor's degree is low.
3) The third independent variable we looked at is the women's share of STEM workers in each state. Studies on the wage differences between STEM and non-STEM jobs have found that STEM workers earn significantly more than non-STEM workers (Socha). Based on this, we argue that states with lower representation of women in STEM fields are likely to have a larger wage gap among genders compared to states with higher representation of women in STEM fields.
4) The fourth independent variable we chose for our model is the percentage of women employed in managerial or professional occupations in each state. As the number of

women who are employed in professional occupations increases, we expect the average income of women to rise. Therefore, we believe the gender wage gap will drop significantly, creating an inverse relationship between these two variables.

The Data:

As mentioned before, our data is comprised of one dependent variable and four independent variables. The dependent variable highlights the topic of our study and through our data we hope to explain the cause(s) of gender wage gap.

The dependent variable labeled “gender wage gap percentage” is an outcome of a calculation we made on the data acquired from a report published by the American Association of University Women (AAUW). This report contains data for the median incomes for both genders at a state level. Since our topic focuses on the gender wage gap, we chose to calculate the wage gap from the income data we collected. Additionally, we chose to express the gap in percentage by using the given formula:

Pay gap% = ([Men’s median income– women’s median income] / Men’s median income)*100

The first independent variable we chose is the “women’s poverty rate” across every state in the U.S., including Washington D.C. We obtained the data for this variable from the National Women’s Law Center. We chose to use the poverty rates of women instead of the poverty rates of men because our dependent variable shows that the wages for women are significantly less than men. Also, May Cohen’s research explains that women with higher poverty rates are likely to end up working in low-paying jobs, thus explaining the wage gap and poverty rate correlation (Cohen 951). This factor displays the adverse effect of poverty on average income of women (See Appendix B.1.).

Our second independent variable expresses the “percentage of women with bachelor’s degree”, in the same survey region used for our first independent variable. We acquired the data for this variable from the National Center for Educational Statistics. We used percentage of women with bachelor’s degree because on average, college graduates have significantly higher wages than non-college graduates (Valleta). Our argument is that women with a bachelor’s degree should earn comparatively similar average wages to men with a bachelor’s degree. If this is the case, we should expect the gender wage gap to be lower (See Appendix B.2.).

The third independent variable we picked measures the “percentage of women among all STEM field workers”. We gathered this data from a report published by the Institute for Women’s Policy Research, under the heading “Status of Women in the States”. We selected STEM fields in particular because the jobs STEM fields have to offer are mostly technical, and require specialized workers for each position. Because of this, the wage tends to be higher for these types of jobs. Additionally, the jobs in STEM fields are given based on skill and not gender, so this should give men and women an equal hiring opportunity. With this data, we are expecting to see an inverse relationship in the percentage of women among STEM field workers and gender wage gap percentage (See Appendix B.3).

Our fourth and final independent variable measures the percentage of “women employed in managerial or professional occupations.” Similar to the third independent variable, we accumulated our data from a report published by the Institute for Women’s Policy Research under the heading “Status of Women in the States”. We specifically selected managerial or professional occupations instead of minimum wage jobs because the representation of women in professional jobs tends to be lower than men. In addition to this, workers in these jobs tend to earn higher wages. Because of this, we will be measuring to see if women in these positions have an effect on the gender wage gap in each state. (See Appendix B.4.).

The Results:

After performing a regression analysis on our data with a 95% confidence level, we ended up with the following equation:

$Y = 35.36 - 0.68(PRatex_1) - 0.9013(BDegreex_2) + 0.046(SWorkersx_3) + 0.518(Occupationx_4) + e$

The numerical values in this equation represent the beta-coefficients. The value for B_0 is 35.36, and this value represents the y-intercept in this model. The remaining beta-coefficients (B_1-B_4) represent the independent variables in this model. The value of B_1 is -0.68, which shows that the there is an inverse correlation between women's poverty rates and the gender wage gap percentage. This result did not meet our expectation for this variable. Likewise, the value for B_2 is -0.9013 shows an inverse correlation between percentage of women with bachelor’s degree and the gender wage gap percentage. Unlike the previous variable, this result managed to meet our expectation, displaying that as the percentage of women with bachelor’s degree increases, the gender wage gap percentage decreases. The value for the percentage of women’s share of STEM workers, represented by B_3 is 0.046. Because this is a positive value it explains that there is a positive correlation between the percentage of women’s share of STEM workers and the gender wage gap. This result is not consistent with our expectation for the variable because the coefficient value points toward the conclusion that if the women’s share of STEM workers increases, the gender wage gap increases as well. The value for B_4 is 0.518, which shows that Percentage of women employed in managerial or professional occupations is positively correlated with the gender wage gap percentage. This is fascinating because it reveals that as the number of women in a managerial or professional occupation increases, the gender wage gap percentage increases, which is the opposite of what we expected.

To determine the statistical significance of the coefficients, we performed t-tests and p-value tests on each independent variable. The result from the t-test on the coefficient of women's poverty rate shows that this variable is statistically significant in determining the gender wage gap percentage (see appendix A.2.). However, the results from our p-value test reveal otherwise (see appendix A.6.). Because of this discrepancy, we are not able to determine if this variable has any significant impact on our model. The t-test result for the percentage of women with a bachelor's degree shows statistical significance (see appendix A.3.). This result is consistent with the result from our p-value test, again displaying statistical significance for this variable (see appendix A.7.).

The results of t-tests on the percentage of women's share in STEM jobs and on the percentage of women employed in professional or managerial occupations, reveal that both of these variables are not statistically significant in determining the gender wage gap percentage (see appendix A.4. and A.5.). This result is further supported by the p-statistic test we conducted on both of these variables because the p-values for these variables also indicate that they are not statistically significant to our dependent variable (see appendix A.8 and A.9).

The adjusted R^2 represents the percentage of variation of the dependent variable explained by the independent variables. Since the R^2 (0.23) is greater than the adjusted R^2 (0.16) we can infer that this decrease is caused by independent variables which have little to no correlation to the dependent variable.

Through the F-Test, we calculated the overall fit of the variables in our model. The F-statistic for our model is 3.54 which is well above our critical value, showing the model is viable in explaining the causes of the gender wage gap percentage (see appendix A.1.).

The multicollinearity test performed on the model indicates that there is a negative correlation between the first independent variable and the second independent variable, with the correlation value being -0.57. Likewise, there is a negative correlation between the first independent variable and the third independent variable for which the correlation value is -0.044. In addition to that, there is also a negative correlation between the first independent variable and the fourth independent variable equating to -0.306. However, the second and the third independent variables show a positive correlation with the value coming out to be at 0.524. The correlation between the third and the fourth independent variables shows a positive correlation on a similar scale with the correlation value of 0.584. The second and the fourth independent variables show the highest positive correlation value of 0.856.

On our heteroskedasticity test we ended up not rejecting the null hypothesis, which shows that there is no heteroskedasticity present in our model (see appendix A.10.). Because of this we can assume constant variance on the error terms. We also performed the Durbin Watson test on our regression model to check for any autocorrelation (see appendix A.11.). The results from the test show positive first order autocorrelation.

Conclusion:

Looking at our results from our F-test, our model shows the trends in the gender wage gap (see appendix A.1.). However, the results from our t-tests and p-value tests show that three of the four variables are not significant in determining the gender wage gap. Through our test results, we saw that there is an inverse relation between the women's poverty rates and the gender wage gap which is not what we expected. However, the results from the statistical tests indicate that women's poverty rates are not statistically significant in determining the gender wage gap. Thus we can claim that the gender wage gap will still be prevalent even if we manage to eliminate or lower the women's poverty rates across the United States. This implies that gender wage gap is not limited to the low income groups but persists regardless of the financial conditions of women. From these findings, we can infer that gender discrimination plays a large role in the gender wage gap.

Likewise, the percentage of women with a bachelor's degree also shows an inverse relationship with the gender wage gap. This claim is further evidenced by the results from the statistical significance tests which show that the percentage of women with a bachelor's degree affect the gender wage gap. So, we can firmly state that as the percentage of women with a bachelor's degree goes up, the gender wage gap decreases. This finding is practical in our studies and the gender wage gap can be reduced if policies are implemented to support the education of women.

The results for the women's share of all STEM workers show a direct relationship with the gender wage gap through the coefficient values. The results indicate that if women's share among the STEM workers increases, than the wage gap increases as well. This does not match with our theoretical assumption. However, this coefficient value is close to zero, suggesting little statistical significance. Additionally, the results of the statistical significance tests show that the percentage of women's share of all STEM workers has no statistical significance in relation to the gender wage gap. This implies that our claim of uniformity in wages for both of the genders across the technical jobs which STEM fields offer is false. From this we can draw that there is gender wage discrimination between men and women in STEM fields, backing up Figart and Kahn's research.

The results for the percentage of women employed in managerial or professional occupations show a similar result to the percentage of women's share in STEM fields. There is a direct relationship between this variable and the gender wage gap. This is not consistent with our theoretical model as we predicted an inverse relationship between these variables. However, the tests for statistical significance for the variable show that percentage of women employed in managerial or professional occupations do not contribute to gender wage gap. This shows that the gender wage gap will be prevalent even if women held a majority of the managerial or professional occupations, implying that gender discrimination is present even in high positions of an organization.

Our research concludes that among the factors such as economic conditions, education status and the types of jobs; education plays a significant role in the gender wage gap. However, we do not have enough evidence to prove that this the only contribution to the gender wage gap. Our research highlights that gender wage discrimination persists in the United States regardless of the economic status and occupation women have. With this, we conclude that our model fails to determine the impact of economic conditions and the job status on the gender wage gap. For this study, we believe that if we incorporated more variables in our model. If we were to conduct this project again we would incorporate studies involve that involve minority groups among the women as well as find data that looks at years of work experience. We believe that this would give a clearer indication of the factors that influence the gender wage gap.

Appendix A - Hypothesis Tests

(A.1.) <u>F-Test on Overall Fit</u>

Step 1: H_0: $B_1 = B_2 = B_3 = B_4 = 0$

H_a: $B_i \neq 0$ ie {1,2,3,4}

Step 2: $\alpha = 5\%$

Reject H_0 if $F^* > F_{cv}$

Numerator = 4, Denominator = 46 (use 40) . . . Use F-table: $F_{cv} = 2.61$

Step 3: $F^* = 3.54$

Step 4: $F^*[3.54] > F_{cv}$ [2.61]

Reject H_0

(A.2.) <u>T-test on Coefficients: Poverty Rates of Women (B1)</u>

Step 1: H_0: $B_1 = 0$

$H_a \neq 0$

Step 2: $\alpha = 5\%$

Reject H_0 if [$t^* > t_{cv}$] or [$t^* < -t_{cv}$]

Degree of freedom = n-(k+1) = 46 (use 40)

Use t-table: $t_{cv} = \pm$ 1.684

Step 3: $t^* = -1.9$

Step 4: $t^*[-1.9] < t_{cv}[-1,684]$

Reject H_0

(A.3.) <u>T-test on Coefficients: Women with Bachelor's Degree (B2)</u>

Step 1: H_0: $B_2 = 0$

$H_a \neq 0$

Step 2: $\alpha = 5\%$

Reject H_0 if $[t^* > t_{cv}]$ or $[t^* < -t_{cv}]$

Degree of freedom = n-(k+1) = 46 (use 40)

Use t table: $t_{cv} = \pm\ 1.684$

Step 3: $t^* = -2.954$

Step 4: $t^*[-2.954] < t_{cv}[-1,684]$

Reject H_0

(A.4.) <u>T-test on Coefficients: % of Women Working in STEM Jobs (B3)</u>

Step 1: H_0: $B_3 = 0$

$H_a \neq 0$

Step 2: $\alpha = 5\%$

Reject H_0 if $[t^* > t_{cv}]$ or $[t^* < -t_{cv}]$

Degree of freedom = n-(k+1) = 46 (use 40)

Use t table: $t_{cv} = \pm\ 1.684$

Step 3: $t^* = 0.167$

Step 4: $t^*[0.167] < t_{cv}[1.684]$

Do not reject H_0

(A.5.) <u>T-test on Coefficients: % of Women Employed in Managerial Occupations (B4)</u>

Step 1: H_0: $B_4 = 0$

$H_a \neq 0$

Step 2: $\alpha = 5\%$

Reject H_0 if [$t^* > t_{cv}$] or [$t^* < -t_{cv}$]

Degree of freedom = n-(k+1) = 46 (use 40)

Use t table: $t_{cv} = \pm 1.684$

Step 3: $t^* = 1.487$

Step 4: $t^*[1.487] < t_{cv}[1684]$

Do not reject H_0

(A.6.) <u>P-value statistics test: Poverty Rates of Women (B1)</u>

Step 1: H_0: $B_1 = 0$

$H_1 \neq 0$

Step 2: $\alpha = 5\%$

Reject H_0 if $P^* < \alpha$ [.05]

Step 3: $P^* = .063$

Step 4: $P^*[.063] > \alpha[.05]$

Do not reject H_0

(A.7.) <u>P-value statistics test: Women with Bachelor's Degree (B2)</u>

Step 1: H_0: $B_2 = 0$

$H_1 \neq 0$

Step 2: $\alpha = 5\%$

Reject H_0 if $P^* < \alpha$ [.05]

Step 3: $P^* = .0049$

Step 4: $P^*[.0049] < \alpha[.05]$

Reject H_0

(A.8.) <u>P-value statistics test: % of Women Working in STEM Jobs (B3)</u>

Step 1: H_0: $B_3 = 0$

$H_1 \neq 0$

Step 2: $\alpha = 5\%$

Reject H_0 if $P^* < \alpha$ [.05]

Step 3: $P^* = .868$

Step 4: $P^*[.868] > \alpha[.05]$

Do not reject H_0

(A.9.) <u>P-value statistics test: Women Employed in Managerial Occupations (B4)</u>

Step 1: H_0: $B_4 = 0$

$H_1 \neq 0$

Step 2: $\alpha = 5\%$

Reject H_0 if $P^* < \alpha$ [.05]

Step 3: $P^* = .143$

Step 4: $P^*[.143] > \alpha[.05]$

Do not reject H_0

(A.10.) <u>Heteroskedasticity</u>

Step 1: H_0: There is no heteroskedasticity

H_a: there is heteroskedasticity

Step 2: Degrees of freedom for numerator and denominator: [n-d-2 (k+1)]/[2]

=15 . . .Use F table (15,15) : $F_{cv} = 2.4034$

Reject H_0 if $F^* > F_{cv}$

Step 3: F^* = SSE Bottom/SSE Top = 0.8343

Step 4: $F^*[0.8343] < F_{cv}[2.4034]$

Do not reject H_0

(A.11.) **<u>Durbin Watson Test</u>**

Step 1: H_0: There is no first order autocorrelation

H_a: There is positive first order autocorrelation

Step 2: Decision Rule

Reject H_0 if $D_w < d_L$

Do not reject H_0 if $D_w > d_u$

Inconclusive if $d_L < D_w < d_u$

K = 4, n = 51 (use 50) . . . DW table: $d_L = 1.38$, $d_u = 1.72$

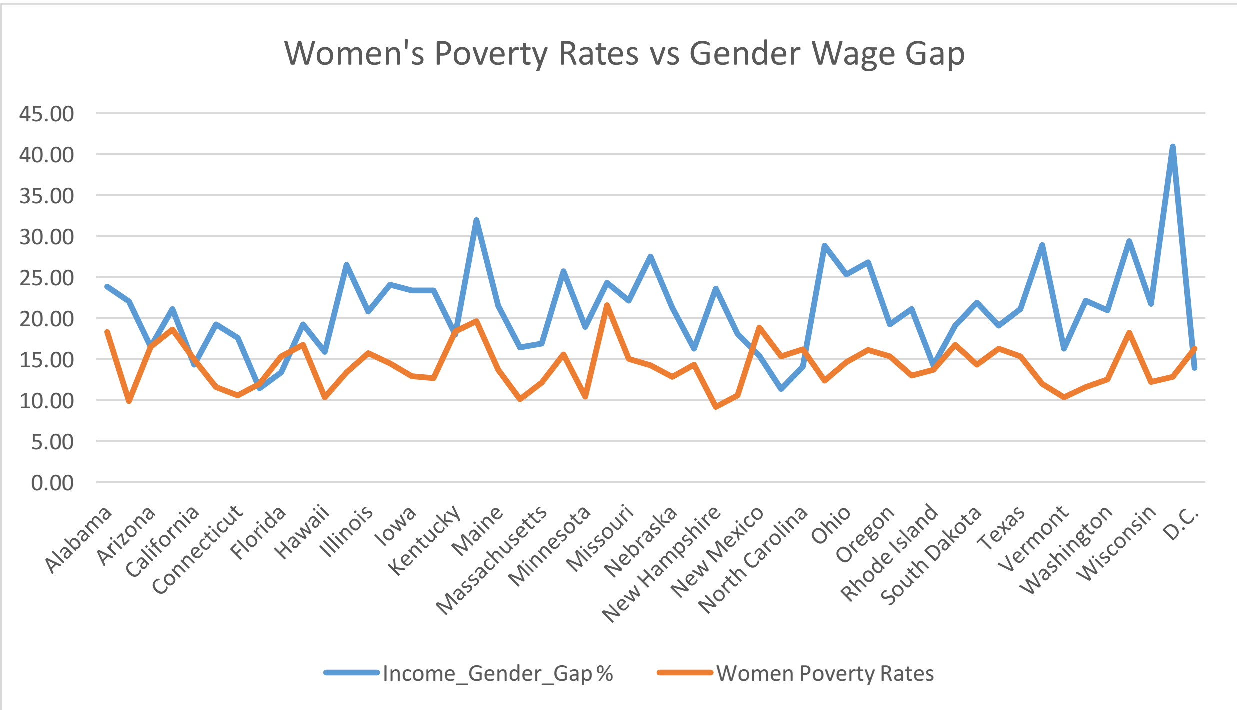
Step 3: $D_w = .00169$

Step 4: $D_w[.00169] < d_L[1.38]$

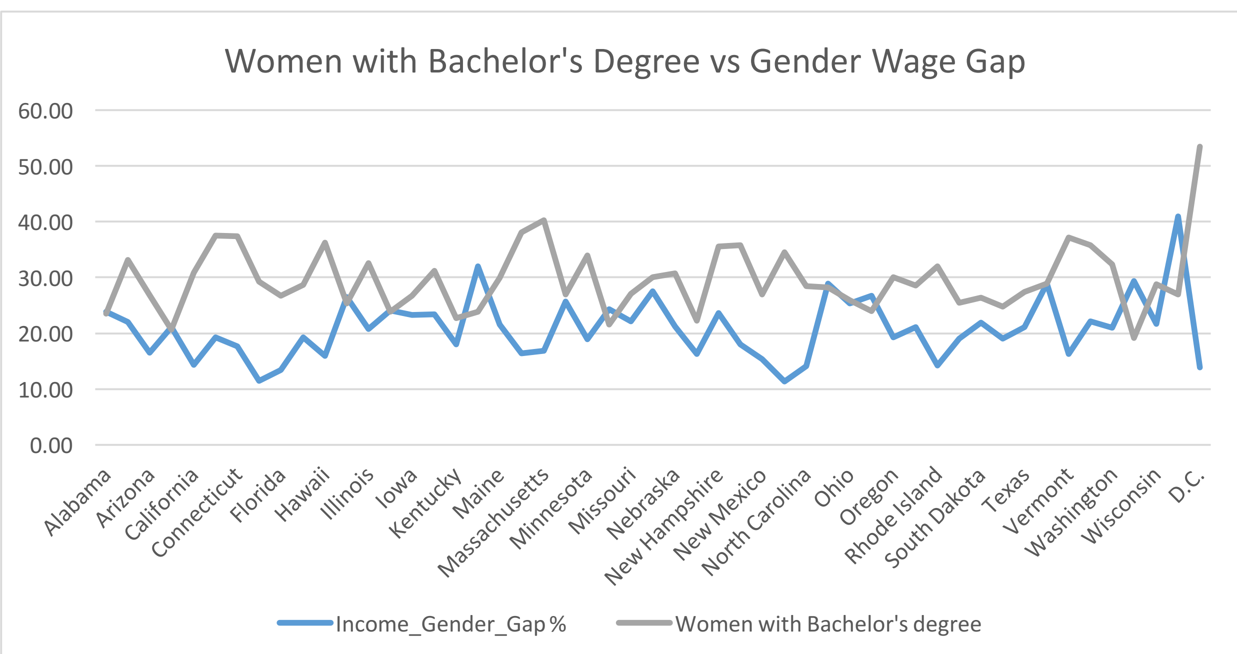
Reject H_0

Appendix B - Graphs

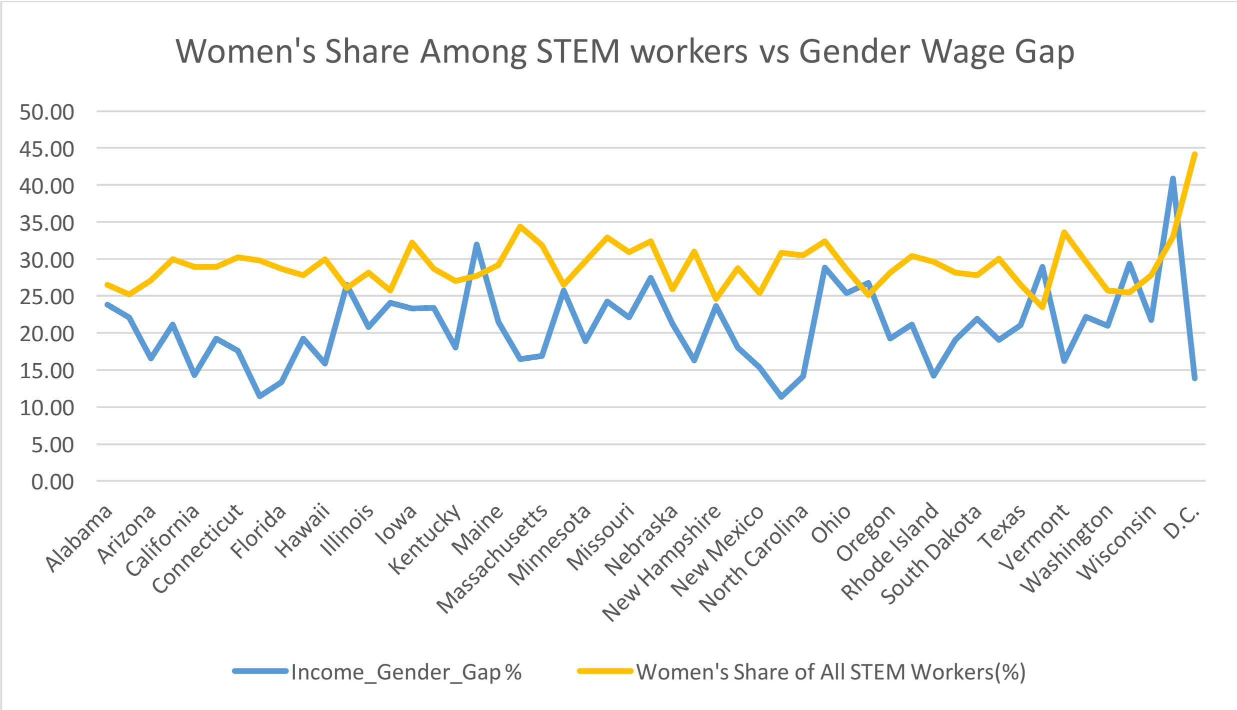
(B.1.)



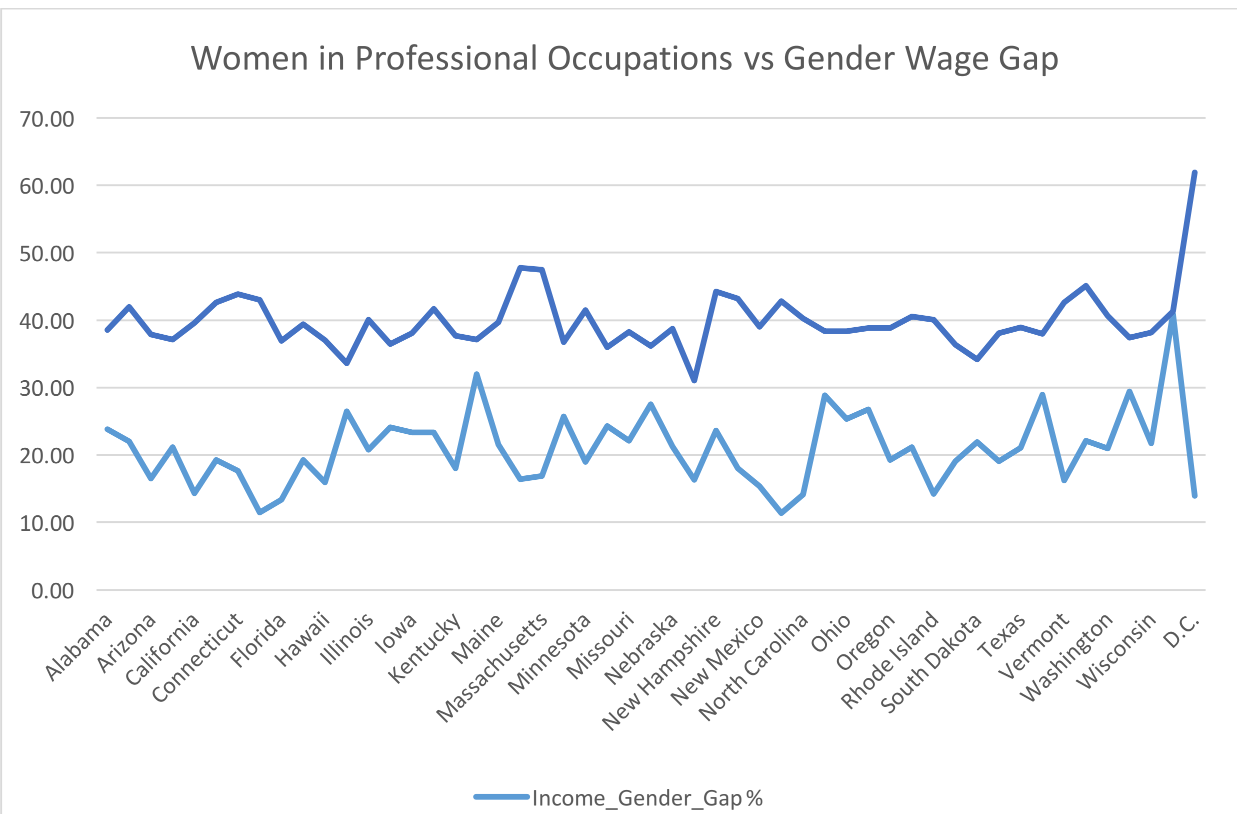
(B.2.)



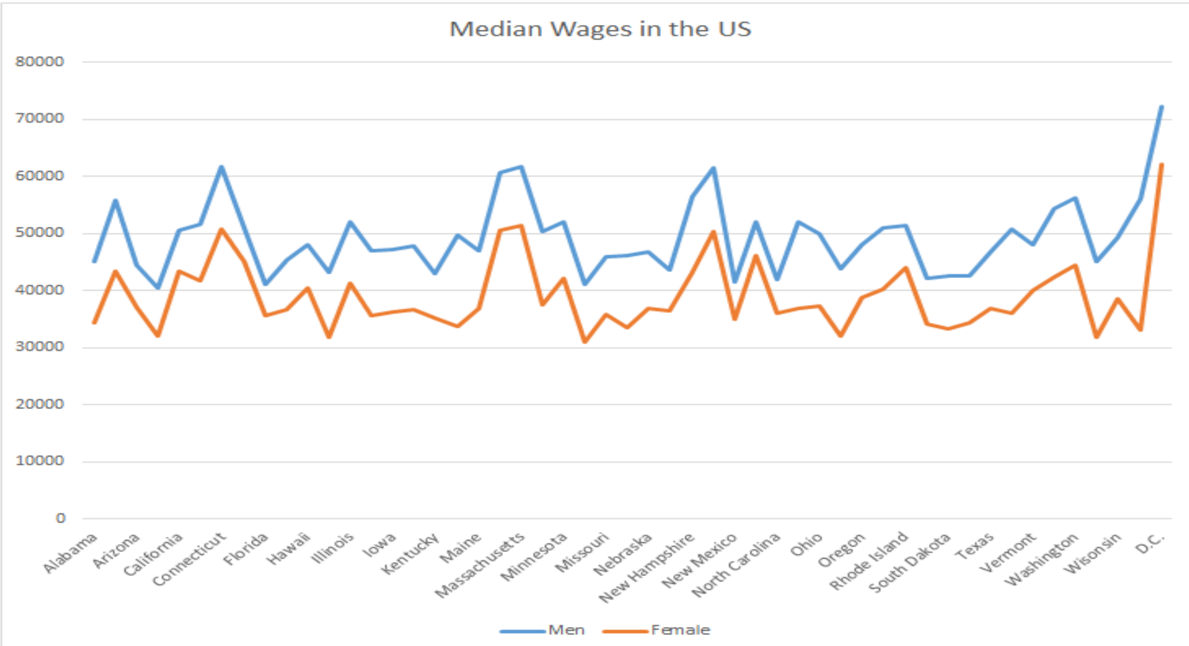
(B.3.)


Women's Share Among STEM workers vs Gender Wage Gap
50.00
45.00
40.00
35.00
30.00
25.00
20.00
15.00
10.00
5.00
0.00
Alabama
Arizona
California
Connecticut
Florida
Hawaii
Illinois
Iowa
Kentucky
Maine
Massachusetts
Minnesota
Missouri
Nebraska
New Hampshire
New Mexico
North Carolina
Ohio
Oregon
Rhode Island
South Dakota
Texas
Vermont
Washington
Wisconsin
D.C.
Income_Gender_Gap %
Women's Share of All STEM Workers(%)

(B.4.)


Women in Professional Occupations vs Gender Wage Gap
70.00
60.00
50.00
40.00
30.00
20.00
10.00
0.00
Alabama
Arizona
California
Connecticut
Florida
Hawaii
Illinois
Iowa
Kentucky
Maine
Massachusetts
Minnesota
Missouri
Nebraska
New Hampshire
New Mexico
North Carolina
Ohio
Oregon
Rhode Island
South Dakota
Texas
Vermont
Washington
Wisconsin
D.C.
Income_Gender_Gap %

(B.5.)


Median Wages in the US
80000
70000
60000
50000
40000
30000
20000
10000
0
Alabama
Arizona
California
Connecticut
Florida
Hawaii
Illinois
Iowa
Kentucky
Maine
Massachusetts
Minnesota
Missouri
Nebraska
New Hampshire
New Mexico
North Carolina
Ohio
Oregon
Rhode Island
South Dakota
Texas
Vermont
Washington
Wisonsin
D.C.
Men
Female

Appendix C – Work in Excel

Works Cited

Cohen, May. 1994. "Impact of Poverty on Women's Health," *Canadian Family Physician*, *40*, 949-957.

McCall, Leslie. 2000. "Gender and the New Inequality: Explaining the College/Non-College Wage Gap," *American Sociological Review,* 65, no. 2, 234-255.

Eswaran, Mukesh. 2014. *Why Gender Matters in Economics*. Princeton University Press, New Jersey.

Deborah M. Figart and Peggy Kahn. 1997. *Contesting the Market: Pay Equity and the Politics ofEconomic Restructuring*, Wayne State University Press, Detroit.

National Center for Education Statistics. (2014). *Baccalaureate and Beyond: A First Look at the Employment Experiences and Lives of College Graduates, 4 Years On*. In T. Socha & . (Eds.). N.p.: U.S. Department of Education.

Valletta, Rob. 2015. "Higher Education, Wages, and Polarization," *Federal Reserve Bank of San Francisco*.

Data Citations

Institute for Women's Policy Research. 2013, B2.1, B2.9. http://statusofwomendata.org/explore-the-data/employment-and-earnings/employment-and-earnings/#TrendsinEmploymentandEarnings

National Center for Education Statistics. 2013, 49. http://nces.ed.gov/pubs2016/2016006.pdf

American Association of University Women. 2016, 8-9.

http://www.aauw.org/aauw_check/pdf_download/show_pdf.php?file=The_Simple-Truth

National Women's Law Center. 2015, 1.

https://nwlc.org/wp-content/uploads/2016/09/Poverty-State-by-State-2016-2.pdf