Senior Project

Department of Economics

Paid Family Leave Policies and Gender Wage Inequality

Camryn Pillar

Spring 2022

Advisor: *Dr. Ali Enami*

Abstract

Over the last few centuries, women's labor force participation rate has grown exponentially. Yet, their problem of inequitable wage and income still exists. Currently women make up 48% of the workforce, but still only make around 20% less than men. Fertility rates and workforce attachment drastically impact wage disparities between men and women. Implementing family leave policies has the potential to combat discrimination towards working mothers. Most of the states rely on the federal mandated unpaid leave, while a select few implements paid leave policies. The anticipation is to find a positive causal relationship with wage equitable economy in states with paid leave policies. Utilizing IPUMS-USA data to determine the effects paid leave policies have on wages using the difference-in-differences technique with a state and year fixed effects. The effects of the paid leave policies on the gender wage gap could not be concluded due to borderline statistically significant results. However, with further investigation, there is statistically significant evidence women benefit from having access to paid leave policies. To further this study, in the future research shall be examined to reveal the true effects paid family leave policies have on the gender wage gap.

Table of Contents

Introduction

The United States of America has striking gender disparities in wage and income. Women's labor participation rate has exponentially progressed in the last century, yet there is still a prevalent wage gap between men and women. Current calculations show that, on average, women earn around eighteen to twenty percent less than men controlling for varying factors such as discrimination, ability, education level, age, and race.[1] One substantial element to incorporate into gender wage research is fertility. Women having children creates a sizable alteration to workforce attachment particularly for the few months postpartum (Baker and Milligan, 2008). Formulating family leave policies can create a more favorable environment to start a family while sustaining a career driven lifestyle (Kluve et al, 2013). Family leave policies vary among states in compensation and time durations allotted. Determining the differential impact, if any, that paid or unpaid leave policies have on the gender wage gap will highlight key correctional elements for combating gender wage inequality. The predicted outcome of this paper is to find that a state's paid leave policy will increase a woman's income compared to states utilizing the federal mandated unpaid time off.

To examine the true effect of paid leave policies on women, data from IPUMS USA is used in this paper to create a difference in difference model with fixed effects. In total there are 8,815,801 individuals observed in the extracted data, 4,324,555 are female and 4,491,246 are male. Each individual is extracted with a range of variables including individual total personal income, wage and salaries, age, race, citizenship status, usual hours worked in a week, educational attainment, employment status, marital status, number of children, occupation,

[1] According to research from the U.S. Census Bureau by Foster, et al. (2020) See: https://www.census.gov/library/working-papers/2020/adrm/CES-WP-20-34.html for the additional information about their methodology.

industry, poverty status, and state of residency for every year. The extracted data for this paper covers years 2000 to 2019 of individuals of 15 to 65 years of age. The six main states of interest, California, Washington, Washington D.C., Rhode Island, New York and New Jersey, will be compared to the immediate surrounding control states. Differing than other research in the field, New York recently passed a family paid leave policy, so it is included in the treatment group.

The economic theory of the labor market can argue why paid leave polices would influence the gender wage gap. Women labor is demanded less than male labor due to discrimination. Men receive a wage premium for having children while women face a wage penalty. Women are the typically the care takers within a household, in which they spend more hours a day doing household cleaning and childcare than men. Employers can think of mothers as distracted on the job, or less productive compared to men. Employers also are hit with the price of hiring and training a new employee if a mother leaves the workforce post pardon, and therefore are more inclines to hire men. Paid leave policies can help with mothers' workforce attachment. Having women more connected to the work force psychologically, then they are less likely to leave the workforce. If paid leave policies help combat this discrimination in the workforce, it can benefit women greatly, improve the gender wage gap.

The predicted outcome of this paper is to find that a state's paid leave policy will increase a woman's income compared to states utilizing the federal mandated unpaid time off. This paper was able to conclude that there is a positive impact on women who have access to the paid leave polices, but further research must be done to determine the impact on the gender wage gap. The rest of the paper is organized as the following: literature review section taking a look at past accredited work done in this field, a data preview section highlighting key facts about the extracted data, theoretical dissection section to further explaining how paid leave policies are connected to the gender wage gap, empirical methodology section illustrating the regression methods of this paper, the results section contain all the findings, and finally the conclusion section summing up this paper.

Literature Review

The historical policy, Family Medical Leave Act (FMLA) of 1993, was the first federal mandate that grants time off after childbirth. Qualified employees will have up to twelve weeks unpaid leave in instances of a new child in the home, caretaking for immediate family members who have fallen ill, the employee having a serious health condition they need to recover from, or if they serve as a military caregiver[2] for an immediate family member. In most states, this is the only policy available, but progress can be seen in few states who have implemented their own paid leave policy. Although these leave options are available for all employees, this doesn't mean the utilization demand of these policies is equal amongst the genders.

These family leave policies, to some extent, create a balance between work and family responsibilities. Historically, women are seen as the homemaker while men are considered the bread winners of the family. Therefore, women complete more of the caretaking tasks in the household compared to men (Gault, et al., 2014). Over the last few centuries, women's time allocation between the household versus the workforce has drastically changed. Pew Research Center (2020), compares gender roles in the family in 1965 to 2011 and finds that women had a statistically significant increase in their hours in the workforce, but men have not escalated their

[2] "Miliray Caregiver" is the next of kin for a service member that is actively missing, seriously injured, or has passed according to the U.S. Department of Labor. See: https://www.dol.gov/agencies/whd/fmla for more information on FMLA policy qualifications

time doing household chores or childcare by the same magnitude[3]. Since women hold a caretaker role, they are more likely to utilize the family leave policies available (Rossin-Slater, 2017). According to Brock (2014), after having children, women tend to stay home full time or go back to work more than men[4]. Therefore, these policies tend to have a greater impact in the women's labor market.

Brock (2014) finds that female labor force participation rate is positively correlated with childcare provision policies. In other words, implementing more policies and programs available that address child care responsibilities actually increases the number of women who decide to join and stay in the workforce. Changes in labor force participation rate have an effect on wages. If women’s labor force participation declines due to bearing children, while men's labor force is not altered, the gender wage gap will become more prominent (Lequien, 2012). Weinstein (2017, p. 592) states, “every 10 percent increase in female labor force participation rates is associated with an increase in real wages of nearly 5 percent” (Weinstein, 2017). This real wage increases also benefits men's salaries as well. Therefore, family leave policies should aim to increase women attachment to the workforce after having children to narrow the gender wage gap. Comparing paid leave policies to unpaid leave policies and their effect on women’s labor force participation can conclude the best form of action to tackle the gender wage gap.

Baum (2003) studied the labor market effects of the FMLA policy and found that it benefited women's labor supply after childbirth. There was a thirteen percentage points

---

[3] Men and women’s time with children and housework did increase which results from more time available in the day due to technology doing tasks that in past years have taken up, such as vacuums, telephones and the internet, etc. This increase was by the same magnitude showing men have not increased their time in caretaking roles compared to women entering the workforce (Pew Research Center, 2020)

[4] 40% of working mothers stated they felt pressure to leave the workforce after childbirth, while only 16% of working fathers felt that same pressure (Brock, 2014)

significant increase in women coming back to their job after giving birth when there are maternity leave policies in place[5]. This research shows FMLA policy positively impacts the gender wage gap, the further question is if this benefit is more or less then a paid policy leave.

In 2003 California became the first state to implement a paid leave program. Employees could qualify for up to six weeks of partial paid leave for time to spend with a newborn or to take time to care for a sick or dying family member (Appelbaum and Milkman, 2011). Researchers Rossin-Slater, Ruhm, and Waldfogel (2013) present that there is robust evidence concluding the program usage in California doubled after this policy was implemented compared to the previous year where FMLA was used. Based on this information, paid leave policies are more beneficial to the gender wage gap because it has a higher influence on women's workforce attachment.

Ample research has been trying to explain why women are more likely to return to work when paid leave policies are implemented compared to unpaid leave. Blau and Lawrence (2017) conclude psychological attributes or noncognitive skills are a leading significant factor in labor participation rates after children. Women account for 46% of the workforce in America, which is almost half, but still face wage inequality (Pew Research Center, 2020). Research shows paid leave positively correlates with a lower rate of postpartum depression and secondary health problems in mothers (Van-Niel et al, 2020). Getting women to feel more connected and supported in the workplace increases the return rates after maternity leave[6].

Cross sectional research on the gender wage gap and family policies shows the gender wage gap connected to women's workforce attachment. This attachment is affected physiologically after a child is born. Policies in which women feel more connected to their job,

[5] Also a seven percent points increase in women looking for a new job after birth with policies present according to Baum (2003)
[6] Marital status is also influential on the rates of return according to Gornick et al. (1998).

have better rates of return after labor. Getting women to stay in the workforce is the one key factor that can help combat gender wage discrimination (Lequien, 2012). There is limited recent studies on this topic. New York recently passed a law almost identical to California's[7]. In my further research I plan to examine paid and unpaid leave laws by state and their impact on the gender wage gap, whilst including New York.

Data Preview

IPUMS USA is the data set that will be utilized in this paper. The IPUMS database provides a range of variables including individual total personal income, wage and salaries, age, race, citizenship status, usual hours worked in a week, educational attainment, employment status, marital status, number of children, occupation, industry, poverty status, and state of residency for every year. The extracted data for this paper covers years 2000 to 2019 of individuals of 15 to 65 years of age. The data includes information from 20 states. The six main states of interest, California, Washington, Washington D.C., Rhode Island, New York and New Jersey, are compared to the immediate surrounding control states. Figure one shows the six

Figure One: *Policy passed and implementation date for states with paid family leave policies*

State	Passed	Implemented
California	2002	July 1, 2004
New Jersey	April 7, 2008	July 1, 2009
New York	April 4, 2016	January 1, 2018
Rhode Island	July 11, 2013	January 1, 2014
Washington	June 30, 2017	October 19, 2017
District of Columbia	2016	April 7, 2017

[7] Bartel et al. (2021) is the leading research

Source: A Better Balance (2022).

main states of interest and the year the paid leave policy was passed and implemented in that state. Implementation date represents the relative year in this paper, discussed more in depth in the empirical methodology section later. California was the first state to pass a paid family leave policy in the United States, which led many others to follow. It is important to note many states are currently in the process of passing similar laws which will not be included in this paper due to lack of data available since they are not implemented yet. The implementation of paid leave policies in the states from figure one, determine what control states and what years are included in the extracted data.

Data from the immediate states is collected as the control group. Figure two shows the state of interest as well as the surrounding states that act as the control group in this paper, as well as the year range that was extracted. The year range was determined based on the implementation date listed in figure one. As noted in the table some of the states of interests are

Figure Two: *Control and Treatment Group's Year Range Data Extraction*

State of Interest	Control States	Year Range
California	Oregon, Nevada, Arizona	2000-2008
New Jersey	New York, Pennsylvania, Delaware, Connecticut, Maryland	2005-2012
New York	Vermont, Massachusetts, Connecticut, New Jersey, Pennsylvania	2015-2019
Rhode Island	Massachusetts, Connecticut, New Hampshire, Vermont	2011-2018
Washington	Oregon, Idaho, Montana	2013-2019
District of Columbia	Maryland, West Virginia, Pennsylvania	2014-2019

Source: Self made

later used as controls for other states of interests. Observations were duplicated into "Silos" so there was no data overlap. This means every state of interest has the control states listed in figure two even if it is also a key state of interest.

In total there are 8,815,801 individuals observed in the extracted data, 4,324,555 are female and 4,491,246 are male. Fissure three shows the averages by gender of some of the

Figure Three: *Nominal Variable Averages by Sex, year 2000 to 2019*

Variable	**Male Average**	**Female Average**
Age	40.85	41.48
Total Personal Income	$53,725.06	$32,707.37
Usual Hours Worked in a Week	33.26	26.18

Source: IPUMS USA-- Own Calculations

nominal variables in the extracted date. The average age of both men and women does not have a huge difference. Men have a much higher total personal yearly income than women, $21,017.69, and typically work more hours a week than women, 7.08 hours. As stated, before the age range was limited from 15-65 based on the definition of "working age." The average age in this data is around 41.48 for female and 40.85 for men. Total personal income includes all financial assets a person gains or loses. In this data the minimum total personal income is negative, and this is because losses are included in this variable. The average total personal income in this data is $32,707.37 for female and $53,725.06 for men. The population includes people of all employment status, which is why the minimum for usual hours worked in a week is zero. Overall, the average hours worked in a week is around 26 hours for females and 33 for males.

Categorical variables were also extracted for all individuals. Figure four shows some of the categorical variables. The majority of the population is white, no children, married, and are

Figure Four: *Categorical Variable Summary Statistics*

Variable	**Category**	**% of Population**
Sex	Male	49.05%
	Female	50.95%
Race	White	75.93%
	Non-White	24.07%

Employment Status	Employed	68.20%
	Unemployed	4.64%
	Not in Labor Force	26.48%
Number of Children	No children	60.20%
	One Child	16.99%
	Two Children	15.02%
	Three or More Children	7.79%

Source: IPUMS USA-- Own Calculations

employed. The population is almost perfectly split between males and females, and therefore is a good representation of the population. All of these categorical variables as well as the nominal variables have an influence on income. These variables will be used in the regression model to determine the true effect paid leave policies have on the gender wage gap.

Theoretical Discussion

The labor market theory and the return on investments in the workforce are the two main economic theories affecting this research. Companies operate at the labor market equilibrium between labor demanded and labor supplied, but equilibriums are not stagnant. Different markets can be created for both men and women since there are some major differences. These differences create the gender wage gap. Men's labor is typically demanded more than female labor. This is in part due to discrimination and historical gender norms and discrimination against women. Higher demand in the labor market means higher wages, and vice versa. Therefore, since mothers are not in demand in the workforce, they actually have lower wages.

There is a competing theory within the labor market on the gender wage gap. Some economists would argue that companies want to make the most profit. Therefore, firms would want to hire more women because they won't have to pay them as much, increasing their profits.

This would increase the demand for female labor and lower the demand of male labor. This would eventually level the wages between men and women to operate at a new equilibrium. This theory is one that states that the existence of the gender wage gap will fix itself. As evident by history, the gender wage gap hasn't made much progress in the last 20th century as would be expected if this theory were true. Therefore, there must be another factor affecting the market, return on investments.

Return on investments for a firm is the fiscal return based on time spent training or educating for a job. More knowledge and experience increase productivity and profitability for forms. Firms aim to be the most profitable possible, therefore predictive, efficient workers are desired. The fiscal responsibility of training and hiring new workers falls on the employers. This means firms are more likely going to hire someone with previous experience which requires less training. Women leaving the workforce for family responsibilities leaves employers the cost of hiring and training new employees. A firm’s cost benefit analysis while hiring a potential job candidate can be altered by discrimination of women, especially mothers. Society's social expectation of women can influence a firm's hiring decisions. Employers are more driven to hire employees that will stay in the workforce longer reducing hiring costs, which in this case is men. Employers can also worry that a woman with children will hinder her work productivity. If a child gets sick or there is a family emergency, it is usually up to the women to handle the situation. Firms could therefore think working women need more time off or will be more distracted at work compared to a man. These discriminatory factor influence the rate in which men are hired over women and created a wage barrier for women.

Family leave policies can aid a mother to feel more connected to the workforce and positively affect the decision to stay in the workforce. Less women leaving the workforce, creates an opportunity for women to advance in the workforce taking on higher positions. Creating a new social expectation that women stay in the workforce after children can combat this underlying discrimination. Firm's not worrying about the fiscal repercussions from women leaving the workforce can encourage a more inclusive work environment for mothers. Paid leave policies if sound to create a stronger workforce attachment for women opposed to an unpaid leave (Lequien, 2012). The question addressed in this paper is whether or not the female benefit of paid family leave policies can fiscally advance them enough to diminish the gender pay gap. The expectation of this research is that paid leave policies will have a greater benefit than unpaid leave policies in the gender wage gap.

Empirical Methodology

There are many different variables that have an effect on income. To examine the true effect different family leave policies, have on gender pay then all of the variables affecting income must be controlled for in the regression model. To better get at the causal effect of the paid family leave on income of women, the difference-in-differences technique is used to compare states with these policies to those who only have federal unpaid family leave. Since states with paid family leave, i.e., treated states, implemented their policies in different years, the difference-in-differences technique uses the concept of relative year as opposed to the calendar year for the time dimension. Specifically, the calendar year in which the policy takes effect in a state is considered "year zero" for that state. The same calendar year is considered year zero for neighboring states of a treated state which constitute the control states (for that specific treated state). The calendar years before "year zero" are identified with a negative value (e.g., -1, -2, -3) and the years after with positive values (e.g., +1, +2, +3). For each treated year, four pretreatment years are included in the analysis, or based on availability. The post treatment years included in the analysis are between 2000 and 2019 depending on the availability of the data. By using this specific quasi-experimental approach, the true effect of a paid leave policy can be found by comparing the changes in outcomes over time between a population that has access to paid leave policies and a population that only has access to FMLA. The regression model that will be used in this study is followed below. The independent variable

Regression One: $\boldsymbol{In_{ist} = B_0 + B_1 Female + B_2 Policy + B_3(Female*Policy) + Year*State + X_{ist} + U_{ist}}$

in regression one is the total personal income on an individual level by state and year. B_0 represents the intercept of the prediction equation. Total income was used as the dependent variable opposed to a variable of salary or wages because of self-employment. The indicator variable of employment status only shows if an individual is employed, not employed, or not in the labor force. Since IPUMS data uses census data, the case of self-employment can get tricky. This information is derived from a questionnaire filled out. In some instances, individuals are unemployed but are still making a wage or salary. As well as some induvial are marked as employed but have no salary or wage indicated. This is because this paper did not have access to self-employment status. Most of the self-employed individuals are men in America. With these incontinences, using wage and salary might not have shown true effects. Therefore, this research uses total personal income as the dependent variable. The coeffect B_1 shows the income discrimination for the indicator variable Female. The variable Policy is also an indicator variable to demonstrate if the individual has access to paid family leave policies available at a certain point in time. If an individual resides in a state with an implemented paid leave policy the policy is equal to one. When policy is equal to zero that means the individual either is in a state that only utilizes FMLA policy, or their state has not implemented a paid leave policy yet. Using both indicator variables the interaction variable Female*Policy shows the true effects of the paid leave policy has on female income. This is the variable of interest in this paper and will be used to test the hypothesis that paid leave policies will have a positive impact on female wages and can combat gender wage inequality. The next interactive term, year*state, represents the year by state fixed effects the model will include. The variable X_{ist} is a placeholder to represent all the variables that will be included in the model as a control. These control variables include age, race, citizenship status, usual hours worked in a week, educational attainment, employment status, marital status, number of children, occupation, industry, and poverty status. The control variables occupation and industry were simplified down based on similarities. See appendix one for the breakdown of these categorical variables. The variable U_{ist} represents the error term of the equation. All of the variables that affect wage are included in this model, therefore the true effect of income for women will be shown in the results. Omitted variable bias is addressing this model with all of the controls in the regression that make the error term less correlated with the variable of interest creating a causal effect. Looking at the true effect on females, the equation simplifies down to the following. This leads us to the following hypothesis:

Null: $B_3 = 0$

Alternative: $B_3 > 0$

If the coefficient policy (B_3) equals zero, then it shows paid leave policies have no effect on women. This paper predicts that implementing Paid leave policies will help the gender wage gap. For this reason, if the coefficient B_3 is greater than one, the true effect for women would increase their income. This increase would show a decrease in the gender wage gap meaning women would make more.

Results

The predicted outcome of this paper is to conclude a state's paid leave policy will increase a woman's income compared to states utilizing the federal mandated unpaid time off. Figure five shows the summarized regression one model. Refer to the empirical methodology

Figure Five: *Regression One*

Variable	**Parameter Estimate**
Female*Policy	-2232.31*** *(87.91)*
Female	-15605.56*** *(43.58)*
Policy	1831.44*** *(627.56)*
State and Year Fixed Effects	**Yes**
Control Variables	**Yes**
F-Statistic	**15131.7**
Adjusted R-Square	**0.3425**
Observations	**8,815,801**
" * " Indicates 10% Significance, " ** " indicates 5% Significance, and " *** "Indicates 1% Significance. Robust standard error is noted under parameter estimate in parenthesis. See Appendix Two for full regression results.	

Source*: IPUMS USA-- Own Calculations*

section of paper for the breakdown and makeup of the model. The key variable of interest, Female*Policy, in regression one is shown as a negative parameter estimate. These results show that paid family level polices has a negative effect on women compared to men. This contradicts research and past literature on this topic. When creating an interactive variable with one positive and one negative with large errors the results may get skewed. This is not an inconsistency that can be resolved without changing the model.

To impact the gender wage gap, paid leave policies need to have a greater positive impact on females than one male. Regression one not aligning with much prior research, a further look into the data can be done. To understand the impact paid leave policies, have on females alone, regression two was created. Figure six shows a summary of regression two which only includes the female population. Since there are no males in this new population the variable Policy will

Figure Six*: Regression Two*

Variable	**Parameter Estimate**
Policy	609.67*** *(90.17)*
Control Variables	**Yes**
State and Year Fixed Effects	**Yes**
Observations	**4,491,246**
Adjusted R-Square	**0.3558**
F-Statistic	**21044.0**
" * " Indicates 10% Significance, " ** " indicates 5% Significance, and " *** "Indicates 1% Significance. Robust standard error is noted under parameter estimate in parenthesis. See Appendix Three for full regression results.	

Source: *IPUMS USA-- Own Calculations*

show the effect paid leave polices impact females compared to females without. These results show a statically significant result, women with paid leave polices make about $609.67 more than women that only have access to unpaid time off. This concludes females are benefited from paid leave policies, but without adding males back into the regression there is no way to determine the impact on the gender wage gap.

To further look into the data, regression one outlines in the empirical section ran on a state level. This will show the effect paid leave policies have on females in each state. Figure seven shows the results of the six regressions. These results show that the paid leave policies

Figure Seven: *Regression One by State*

Variable	**CA**	**DC**	**NJ**	**NY**	**RI**	**WA**
Female*Policy	-2596.25*** *(132.35)*	1157.24*** *(1329.82)*	-4522.74*** *(216.34)*	4575.86*** *(217.90)*	4466.13*** *(536.03)*	-6729.22*** *(327.66)*
Female	-13075.816*** *(97.69)*	-15844.65*** *(110.27)*	-15733.25*** *(67.71)*	-15759*** *(120.13)*	-17650.85*** *(160.77)*	-14840.0*** *(141.58)*
Policy	2855.60*** *(440.98)*	1417.68*** *(195.04)*	1414.27*** *(424.65)*	4726.69*** *(1004.42)*	-3009.814** *(1278.83)*	5552.68*** *(575.51)*
State by Year Fixed Effects	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
Control	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
F-Statistic	**9142.71**	**7691.37**	**13484.60**	**9454.35**	**3838.10**	**3929.00**
Adjusted R-Squared	**0.357**	**0.369**	**0.339**	**0.330**	**0.339**	**0.367**
Observations	**1,646,720**	**1,151,404**	**2,972,079**	**1,703,108**	**717,078**	**625,412**

" * " Indicates 10% Significance, " ** " indicates 5% Significance, and " *** "Indicates 1% Significance. Robust standard error is noted under parameter estimate in parenthesis. See Appendix Full regression available on request.

Source: *IPUMS USA-- Own Calculations*

have a different effect across states. Therefore, when running regression one with all the states, noted in figure five, does not give a clear representation of the impact paid leave policies have one female. Rerunning the same 6 regressions in figure seven, but only using the female population[8] shows that women with access to paid leave policies are better off. Breaking down the regression by state there are still inconsistencies with the coeffect of Policy*Female, as well as a problem with the size of the robust errors being on the larger side. However, only looking at the female population they are better off with paid leave policies. This is reflective of the results shown in figure five and figure six.

[8] Results of these six regressions are available upon request. For all states the coefficient for Policy was greater than zero and was statistically significant.

The difference in difference method, used in all regressions in this paper, is used to estimate the treatment effect by estimating the average treatment effect of the causal effect in the population. This technique of using averages may show reasoning for the inconsistences in the results. Averages may not be the best idea given that policies seem to be behaving differently and in the future each state should be potentially analyzed independently. One cannot guess the direction of the effect on the gender wage gap based on the existing conflicting results from other states.

The inconsistencies of these results can also be explained by employer policies. Even though paid leave policies were passed in the treatment states does mean that women don't have access to similar policies that reside in the control state. Employers have been known to implement policies on a company level to inceptives more individuals to work for them over compactors. This can lead to the indication variable Policy to not fully cover those individuals who have leave policies through their employers. Therefore, the results shown in the regression results could be skewed by this bias.

For future implications on this paper, a regression must be flushed out to determine the impact paid leave policies have on the gender wage gap. Variables sometimes effect males and females differently, such as number of children. What is called the father premium and the mother penalty just proves that one variable can affect males and females differently. Women who have children typically receive less compensation, while males who have children are typically rewarded with more compensation. Because of the difference, the regression one model cannot predict the most accurate effect paid leave polices have on the gender wage gap. In the future, a regression should be done with a state by year by gender fixed effects.

Conclusion

The United States of America has striking gender disparities in wage and income. Women's labor participation rate has exponentially progressed in the last century, yet there is still a prevalent wage gap between men and women. Current calculations show that, on average, women earn around eighteen to twenty percent less than men. Fertility and children have a major impact on income for both men in women. Women who have children typically receive less compensation, while males who have children are typically rewarded with more compensation. Discrimination on the workforce is the leading cause of the gender wage gap. Women's labor is demanded less than men's labor creating a divide in wages. Historically, women are the caretakers of the homes. On average mothers spend more time cleaning the household and caring for the children compared to the fathers, even when they are both equally employed. Therefore, women are expected to stay home after a child if born. Women leaving the workforce for family responsibilities leaves employers the cost of hiring and training new employees. Therefore, a firm is more likely to hire an expecting father compared to an expecting mother.

Combatting this discrimination is key in improving the gender wage gap in America. This paper analyses the effect paid leave policies within states have on the gender wage gap. This paper anticipation was to find a causation that paid leave policies improved the gender wage gap. Using a difference in differences regression technique with sufficient data extracted from IPUMS USA different models were created. This paper concludes females are benefited from paid leave policies but was unable to determine the impact on the gender wage gap. Variables effect each gender differently, number of children for example. Future research must be done to determine the effect paid leave policies have on the gender wage gap. Creating a new regression with a state by year by gender fixed effects should be created. If applicable, if data can be collected on all the individuals' employers policy would make the results more accurate. Even so, without knowing the impact paid leave policies have on gender wage, since there is a positive impact for women, more states should implement paid leave policies.

References

A Better Balance (2022). "Comparative Chart of Paid Family and Medical Leave Laws in the United States." https://www.abetterbalance.org/resources/paid-family-leave-laws-chart/

Appelbaum, E. and R. Milkman (2011). Paid Family Leave Pays Off in California. *Harvard Business Review.*

Anderson, D. J., Binder, M., and Krause, K. (2002). The Motherhood Wage Penalty: Which Mothers Pay It and Why? *The American Economic Review*, *92*(2), 354–358.

Baker, M. and K. Milligan (2008). How Does Job-Protected Maternity Leave Affect Mothers' Employment? *Journal of Labor Economics* 26 (4), 655–691.

Bartel, A. Rossin-Slater, M. Ruhm, C. Slopen, M. and Waldfogel, J. (2021). National Bureau of Economic Research. Working Paper 28672

Baum, C. L. (2003). The Effects of Maternity Leave Legislation on Mothers' Labor Supply after Childbirth. *Southern Economic Journal* 69 (694), 772–799.

Blau, Francine D., and Lawrence M. Kahn. (2017). "The Gender Wage Gap: Extent, Trends, and Explanations." *Journal of Economic Literature*, 55 (3): 789-865.

Borck, R. (2014). *Adieu Rabenmutter*—culture, fertility, female labor supply, the gender wage gap and childcare. *J Popul Econ* 27, 739–765 .

Byker, Tanya S. (2016). "Paid Parental Leave Laws in the United States: Does Short-Duration Leave Affect Women's Labor-Force Attachment?" *American Economic Review*, 106 (5): 242-46.

Gornick, J. C., Meyers, M. K., & Ross, K. E. (1998). Public Policies and the Employment of Mothers: A Cross-National Study. *Social Science Quarterly*, *79*(1), 35–54. http://www.jstor.org/stable/42863766

Kluve, J., M. Tamm, J. Kluve, and M. Tamm (2013). Parental leave regulations, mothers' labor force attachment and fathers' childcare involvement: evidence from a natural experiment. *Journal of Population Economics* 26 (3), 983–1005.

Lequien, L. (2012). The Impact of Parental Leave Duration on Later Wages. *Annals of Economics and Statistics 107/108 (July/December), 267–285.*

"Modern Parenthood." *Pew Research Center's Social & Demographic Trends Project,* Pew Research Center, 10 Sept. 2020, https://www.pewresearch.org/social-trends/2013/03/14/modern-parenthood-roles-of-moms-and-dads-converge-as-they-balance-work-and-family/#fn-16485-2.

Rossin-Slater, M. (2017). Maternity and Family Leave Policy. Working Paper 23069, *National Bureau of Economic Research.*

Rossin-Slater, M., C. J. Ruhm, and J. Waldfogel (2013). The effects of California's paid family leave program on mothers' leave-taking and subsequent labor market outcomes. *Journal of Policy Analysis and Managemen*t 32 (2), 224–245.

Steven Ruggles, Sarah Flood, Sophia Foster, Ronald Goeken, Jose Pacas, Megan Schouweiler and Matthew Sobek. IPUMS USA: Version 11.0 [dataset]. Minneapolis, MN: IPUMS, 2021. https://doi.org/10.18128/D010.V11.0

Van-Niel, Maureen Sayres MD; Bhatia, Richa MD; Riano, Nicholas S. MAS; de Faria, Ludmila MD; Catapano-Friedman, Lisa MD; Ravven, Simha MD; Weissman, Barbara MD; Nzodom, Carine MD; Alexander, Amy MD; Budde, Kristin MD, MPH; Mangurian, Christina MD, MAS The Impact of Paid Maternity Leave on the Mental and Physical Health of Mothers and Children: A Review of the Literature and Policy Implications, Harvard Review of Psychiatry: 3/4 2020 - Volume 28 - Issue 2 - p 113-126 doi: 10.1097/HRP.0000000000000246

Whitehouse, G. (2005). Policy and Women's Workforce Attachment. *Just Policy: A Journal of Australian Social Policy,* (35), 22–30.

Weinstein, A, (2017). Working women in the city and urban wage growth in the United States. Journal of Regional Science. 57(4). 591-610.

Appendix One

Industry Categorization:

Industry Category	Coding Assignment
Agriculture, Forestry, Fishing, and Hunting, and Mining	1
Construction	2
Manufacturing	3
Wholesome Trade	4
Retail Trade	5
Transportation and Warehousing, and Utilities	6
Information	7
Finance and Insurance, and Real Estate, and Rental and Leasing	8
Professional, scientific, and management, and administrative, and Waste Management Services	9
Educational Services, and Health care and social work	10
Arts, entertainment, and recreation, and accommodation and food services	11
Other service, except public administration	12
Public administration	13
Military	14

Occupation Categorization:

Occupation Category	Code Assignment
Management, Business& financial operation, professional, and related occupations	1
Service occupations	2
Sale and Office Occupations	3
Construction, Extraction, and Maintenance Occupations	4
Production, Transportation, and Material Moving Occupations; Military	5

Appendix Two

Regression One

Variable	Parameter Estimate
Intercept	-36465.36*** *(1117.68)*
Female*Policy	-2231.31*** *(87.91)*
Female	-15605.56*** *(43.57)*
Policy	1831.44*** *(627.56)*
Age	532.62*** *(1.59)*
Poverty	90.63*** *(0.12)*
Usual Weekly Work Hours	912.24*** *(1.46)*
Race	
White	1021.23*** *(394.87)*
Black	-2176.79*** *(398.51)*
Native	-825.72*** *(452.71)*
Chinese	5076.15*** *(414.44)*
Japanese	5544.43*** *(516.70)*
Other Asian or Pacific Islander	1895.25*** *(404.05)*
Other Race	-3351.73*** *(402.90)*
Two Major Races	-715.22* *(412.18)*
Three or More Major Races	.

Marital Status	
Married, Spouse Present	-92.18 *(56.51)*
Married, Spouse Absent	2934.55*** *(131.35)*
Separated	2297.14*** *(130.82)*
Divorced	2454.75*** *(73.77)*
Widowed	5568.60*** *(144.52)*
Never Married/Single	.
Number of Children	
One	-8076.72*** *(1074.13)*
Two	-265.56 *(1074.11)*
Three	4209.06*** *(1075.61)*
Four	4820.24*** *(1082.20)*
Five	4595.15*** *(1106.63)*
Six	5842.85*** *(1168.45)*
Seven	4480.10*** (1288.29)
Eight	3530.29** *(1482.06)*
Nine +	.
Employment Status	
Employed	6300.54*** *(65.05)*
Unemployed	793.84*** *(92.95)*
Not in Labor Force	.

Occupation	
Management, Business& Financial Operation, Professional, and Related	22872.81*** *(110.49)*
Service Occupations	5270.49*** *(125.60)*
Sale and Office Occupations	555.07*** *(133.62)*
Construction, Extraction, and Maintenance Occupations	2812.88*** *(150.20)*
Production, Transportation, and Material Moving Occupations; Military	.
Industry	
Agriculture, Forestry, Fishing, and Hunting, and Mining	-13986.36*** *(316.60)*
Construction	-4806.73*** *(282.36)*
Manufacturing	4343.63*** *(279.04)*
Wholesome Trade	5674.74*** *(296.01)*
Retail Trade	-812.40*** *(278.28)*
Transportation and Warehousing, and Utilities	-1787.50*** *(286.30)*
Information	14968.38*** *(298.42)*
Finance and Insurance, and Real Estate, and Rental and Leasing	26514.46*** *(282.21)*
Professional, Scientific, Management, Administrative, and Waste Management Services	16202.09*** *(277.83)*
Educational Services, and Health care and social work	4888.20*** *(275.84)*
Arts, entertainment, and recreation, food services	-4100.48*** *(279.89)*
Other Service, Except Public Administration	-4100.48*** *(286.42)*
Public Administration	8067.72*** *(285.05)*

Military	.
State and Year Fixed Effects	**Yes**
F-Statistic	**15131.7**
Adjusted R-Square	**0.3425**
Observations	**8,815,801**

“ * ” Indicates 10% Significance, “ ** ” indicates 5% Significance, and “ *** ”Indicates 1% Significance. Robust standard error is noted under parameter estimate in parenthesis. Reference group noted by “.” in table.

Source: *IPUMS USA-- Own Calculations*

Appendix Three

Regression Two

Variable	**Parameter Estimate**
Intercept	-27309.30*** *(802.92)*
Policy	609.67*** *(90.17)*
Age	373.20*** *(1.60)*
Poverty	69.47*** *(0.13)*
Usual Weekly Work Hours	883.19*** *(1.52)*
Race	
White	421.05 *(384.77)*
Black	-935.67** *(388.42)*
Native	-788.69* *(444.66)*
Chinese	4817.96*** *(404.00)*
Japanese	945.40* *(496.37)*
Other Asian or Pacific Islander	1602.63*** *(394.01)*
Other Race	-2328.40*** *(393.18)*
Two Major Races	-883.82** *(403.53)*
Three or More Major Races	.
Marital Status	

Married, Spouse Present	-5499.65*** *(56.27)*
Married, Spouse Absent	1227.31*** *(137.79)*
Separated	1082.87*** *(123.79)*
Divorced	3565.83*** *(72.28)*
Widowed	4564.29*** *(122.06)*
Never Married/Single	.
Number of Children	
One	-5743.60*** *(1074.60)*
Two	-5136.56*** *(1074.91)*
Three	-860.95 *(1074.92)*
Four	-367.20 *(1076.32)*
Five	232.79 *(1082.51)*
Six	972.84 *(1105.42)*
Seven	2289.37* *(1282.14)*
Eight	2711.30* *(1476.21)*
Nine +	.
Employment Status	
Employed	1359.17*** *(71.60)*
Unemployed	-2939.18*** *(95.86)*

Not in Labor Force	.
Occupation	
Management, Business& Financial Operation, Professional, and Related	-7113.94*** *(115.93)*
Service Occupations	17037.31*** *(125.52)*
Sale and Office Occupations	2200.24*** *(139.06)*
Construction, Extraction, and Maintenance Occupations	6203.82*** *(182.06)*
Production, Transportation, and Material Moving Occupations; Military	.
Industry	
Agriculture, Forestry, Fishing, and Hunting, and Mining	4762.80*** *(495.72)*
Construction	-13472.86*** *(542.38)*
Manufacturing	-3321.50** *(591.75)*
Wholesome Trade	1050.38** *(497.43)*
Retail Trade	1214.56** *(512.19)*
Transportation and Warehousing, and Utilities	-6556.42*** *(494.53)*
Information	-3561.12*** *(507.02)*
Finance and Insurance, and Real Estate, and Rental and Leasing	6996.24*** *(508.48)*
Professional, Scientific, Management, Administrative, and Waste Management Services	7398.89*** *(496.31)*
Educational Services, and Health care and social work	5510.18*** *(494.68)*
Arts, entertainment, and recreation, and accommodation and food services	-2489.23*** *(492.28)*

Other Service, Except Public Administration	-8262.51*** *(498.46)*
Public Administration	4783.80*** *(499.27)*
Military	.
State and Year Fixed Effects	**Yes**
F-Statistic	**21044.0**
Adjusted R-Square	**0.356**
Observations	**4491246**

“ * ” Indicates 10% Significance, “ ** ” indicates 5% Significance, and “ *** ”Indicates 1% Significance. Robust standard error is noted under parameter estimate in parenthesis. Reference group noted by “.” in table.

Source: *IPUMS USA-- Own Calculations*

Appendix Four

SAS Academics Code:

```
LIbname econ "/home/u53962797/Semester.Project";

PROC IMPORT
DATAFILE="/home/u53962797/Semester.Project/usa_00084.csv"
OUT= CA1
DBMS=csv
  REPLACE;
RUN;quit;
PROC IMPORT
DATAFILE="/home/u53962797/Semester.Project/usa_00085.csv"
OUT= OR1
DBMS=csv
  REPLACE;
RUN;quit;
PROC IMPORT
DATAFILE="/home/u53962797/Semester.Project/usa_00086.csv"
OUT= NV1
DBMS=csv
  REPLACE;
RUN;quit;
PROC IMPORT
DATAFILE="/home/u53962797/Semester.Project/usa_00087.csv"
OUT= AZ1
DBMS=csv
  REPLACE;
RUN;quit;
/*PROC IMPORT
DATAFILE="/home/u53962797/Semester.Project/usa_00088.csv"
OUT= UT1
DBMS=csv
  REPLACE;
RUN;quit;
PROC IMPORT
DATAFILE="/home/u53962797/Semester.Project/usa_00089.csv"
OUT= ID1
DBMS=csv
  REPLACE;
RUN;quit;*/
Data CA_Data;
set CA1 OR1 NV1 AZ1 /*UT1 ID1*/;
if Incwage=999999 then delete;
if incwage =999998 then delete;
if age < 15 then delete;
if age > 65 then delete;
ryear=year-2004;
silo= "CA";
run;quit;



PROC IMPORT
```

```
DATAFILE="/home/u53962797/Semester.Project/usa_00092.csv"
OUT= WA2
DBMS=csv
  REPLACE;
RUN;quit;
PROC IMPORT
DATAFILE="/home/u53962797/Semester.Project/usa_00093.csv"
OUT= OR2
DBMS=csv
  REPLACE;
RUN;quit;
PROC IMPORT
DATAFILE="/home/u53962797/Semester.Project/usa_00094.csv"
OUT= ID2
DBMS=csv
  REPLACE;
RUN;quit;
PROC IMPORT
DATAFILE="/home/u53962797/Semester.Project/usa_00095.csv"
OUT= MT2
DBMS=csv
  REPLACE;
RUN;quit;
/*PROC IMPORT
DATAFILE="/home/u53962797/Semester.Project/usa_00096.csv"
OUT= NV2
DBMS=csv
 REPLACE;
RUN;quit;
PROC IMPORT
DATAFILE="/home/u53962797/Semester.Project/usa_00097.csv"
OUT= WY2
DBMS=csv
  REPLACE;
RUN;quit;*/
Data WA_Data;
set WA2 OR2 ID2 MT2 /*NV2 WY28*/;
if Incwage=999999 then delete;
if incwage =999998 then delete;
if age < 15 then delete;
if age > 65 then delete;
ryear=year-2017;
silo="WA";
run; quit;


PROC IMPORT
DATAFILE="/home/u53962797/Semester.Project/usa_00098.csv"
OUT= DC3
DBMS=csv
  REPLACE;
RUN;quit;
PROC IMPORT
DATAFILE="/home/u53962797/Semester.Project/usa_00099.csv"
OUT= MD3
DBMS=csv
```

```
  REPLACE;
RUN;quit;
PROC IMPORT
DATAFILE="/home/u53962797/Semester.Project/usa_00100.csv"
OUT= WV3
DBMS=csv
  REPLACE;
RUN;quit;
PROC IMPORT
DATAFILE="/home/u53962797/Semester.Project/usa_00101.csv"
OUT= VA3
DBMS=csv
  REPLACE;
RUN;quit;
PROC IMPORT
DATAFILE="/home/u53962797/Semester.Project/usa_00102.csv"
OUT= PA3
DBMS=csv
  REPLACE;
RUN;quit;
Data DC_Data;
set DC3 MD3 WV3 VA3 PA3;
if Incwage=999999 then delete;
if incwage =999998 then delete;
if age < 15 then delete;
if age > 65 then delete;
ryear= year-2017;
Silo="DC";
run;quit;

PROC IMPORT
DATAFILE="/home/u53962797/Semester.Project/usa_00103.csv"
OUT= RI4
DBMS=csv
  REPLACE;
RUN;quit;
PROC IMPORT
DATAFILE="/home/u53962797/Semester.Project/usa_00104.csv"
OUT= MA4
DBMS=csv
  REPLACE;
RUN;quit;
PROC IMPORT
DATAFILE="/home/u53962797/Semester.Project/usa_00105.csv"
OUT= CT4
DBMS=csv
  REPLACE;
RUN;quit;
PROC IMPORT
DATAFILE="/home/u53962797/Semester.Project/usa_00106.csv"
OUT= NH4
DBMS=csv
  REPLACE;
RUN;quit;
PROC IMPORT
DATAFILE="/home/u53962797/Semester.Project/usa_00107.csv"
```

```
OUT= VT4
DBMS=csv
  REPLACE;
RUN;quit;
Data RI_Data;
set RI4 MA4 CT4 NH4 VT4;
if Incwage=999999 then delete;
if incwage =999998 then delete;
if age < 15 then delete;
if age > 65 then delete;
ryear=year-2014;
silo="RI";
run; quit;

PROC IMPORT
DATAFILE="/home/u53962797/Semester.Project/usa_00108.csv"
OUT= NY5
DBMS=csv
  REPLACE;
RUN;quit;
PROC IMPORT
DATAFILE="/home/u53962797/Semester.Project/usa_00109.csv"
OUT= VT5
DBMS=csv
  REPLACE;
RUN;quit;
PROC IMPORT
DATAFILE="/home/u53962797/Semester.Project/usa_00110.csv"
OUT= MA5
DBMS=csv
  REPLACE;
RUN;quit;
PROC IMPORT
DATAFILE="/home/u53962797/Semester.Project/usa_00111.csv"
OUT= CT5
DBMS=csv
  REPLACE;
RUN;quit;
PROC IMPORT
DATAFILE="/home/u53962797/Semester.Project/usa_00112.csv"
OUT= NJ5
DBMS=csv
  REPLACE;
RUN;quit;
PROC IMPORT
DATAFILE="/home/u53962797/Semester.Project/usa_00113.csv"
OUT= PA5
DBMS=csv
  REPLACE;
RUN;quit;
Data NY_Data;
set NY5 VT5 MA5 CT5 NJ5 PA5;
if Incwage=999999 then delete;
if incwage =999998 then delete;
if age < 15 then delete;
if age > 65 then delete;
```

```
ryear=year-2018;
silo="NY";
run;quit;

PROC IMPORT
DATAFILE="/home/u53962797/Semester.Project/usa_00114.csv"
OUT= NJ6
DBMS=csv
  REPLACE;
RUN;quit;
PROC IMPORT
DATAFILE="/home/u53962797/Semester.Project/usa_00115.csv"
OUT= NY6
DBMS=csv
  REPLACE;
RUN;quit;
PROC IMPORT
DATAFILE="/home/u53962797/Semester.Project/usa_00116.csv"
OUT= PA6
DBMS=csv
  REPLACE;
RUN;quit;
PROC IMPORT
DATAFILE="/home/u53962797/Semester.Project/usa_00117.csv"
OUT= DE6
DBMS=csv
  REPLACE;
RUN;quit;
PROC IMPORT
DATAFILE="/home/u53962797/Semester.Project/usa_00118.csv"
OUT=CT6
DBMS=csv
  REPLACE;
RUN;quit;
PROC IMPORT
DATAFILE="/home/u53962797/Semester.Project/usa_00119.csv"
OUT= MD6
DBMS=csv
  REPLACE;
RUN;quit;
Data NJ_Data;
set NJ6 NY6 PA6 DE6 CT6 MD6;
if Incwage=999999 then delete;
if incwage =999998 then delete;
if age < 15 then delete;
if age > 65 then delete;
ryear=year-2009;
silo="NJ";
run; quit;

Data CA_DataNew(drop=DENSITY METRO CBSERIAL);
        set CA_Data;
        Density2= input(DENSITY, Best5.);
        Metro2 = input(METRO, Best5.);
        CBSERIAL2 = input(CBSERIAL, Best5.);
run;
```

```
Data Data;
set CA_DataNew WA_Data DC_Data RI_Data NY_Data NJ_Data;
if statefip=6 and year >= 2004 then Policy =1;
else if statefip=34 and year >=2009 then Policy=1;
else if statefip = 36 and year>= 2018 then policy=1;
else if statefip= 44 and year >= 2014 then policy = 1;
else if statefip= 53 and year>=2017 then policy =1;
else if statefip=11 and year >= 2017 then policy=1;
else Policy=0;
if sex= 2 then female=1;
else female=0;
femalepolicy=female*policy;
if year = 2020 then delete;
run; quit;

proc freq data=data;
table statefip*policy;
run; quit;

Data Data;/*categorization of ind*/
set Data;
if year >=2003 and ind >=170 and ind<=490 then ind2=1;
        else if year>=2003 and ind = 770 then ind2=2;
        else if year >=2003 and ind >=1070 and ind <=3990 then ind2=3;
        else if year >=2003 and ind >= 4070 and ind <=4590 then ind2=4;
        else if year >=2003 and ind >= 4670 and ind <= 5790 then ind2=5;
        else if year >=2003 and ind >= 6070 and ind <=6390 then ind2=6;
        else if year >=2003 and ind >= 570 and ind <= 690 then ind2=6;
        else if year >=2003 and ind >= 6470 and ind <=6780 then ind2=7;
        else if year >=2003 and ind >= 6870 and ind <= 7190 then ind2=8;
        else if year >=2003 and ind >=7270 and ind <= 7790 then ind2=9;
        else if year >=2003 and ind >= 7860 and ind<= 8470 then ind2=10;
        else if year >=2003 and ind >= 8561 and ind <= 8690 then ind2=11;
        else if year >=2003 and ind >= 8770 and ind<=9290 then ind2=12;
        else if year >=2003 and ind >=9370 and ind<=9590 then ind2=13;
        else if year >=2003 and ind>= 9670 and ind<= 9870 then ind2=14;
        else if Year >=2003 then ind2=0;
If year<=2002 and Ind= 77 then ind2=2;
        else if year<=2002 and Ind>=17 and ind<=399 then ind2 =3;
        else if year<=2002 and Ind>=407 and ind<=459 then ind2=4;
        else if year<=2002 and Ind>=467 and ind<=579 then ind2=5;
        else if year<=2002 and Ind>=607 and ind<=69 then ind2 =6;
        else if year<=2002 and Ind>=647 and ind<=679 then ind2=7;
        else if year<=2002 and Ind>=687 and ind<=719 then ind2=8;
        else if year<=2002 and Ind>=727 and ind<=779 then ind2 = 9;
        else if year<=2002 and Ind>=786 and ind<=847 then ind2=10;
        else if year<=2002 and Ind>=856 and ind<=869 then ind2=11;
        else if year<=2002 and Ind>=877 and ind<=929 then ind2=12;
        else if year<=2002 and Ind>=937 and ind<=959 then ind2 = 13;
        else if year<=2002 and Ind>=967 and ind<=987 then ind2=14;
        else if year<=2002 then ind2=0;
run; quit;

data Data;/*Occupation categorization*/
set Data;
```

```
if year<=2017 and occ=>1 and occ=<95 then occ2=1;
        else if year<=2017 and occ>=100 and occ<=354 then occ2=1;
        else if year<=2017 and occ>=360 and occ<=465 then occ2=2;
        else if year<=2017 and occ>=470 and occ<=613 then occ2=3;
        else if year<=2017 and occ>=620 and occ<=762 then occ2=4;
        else if year<=2017 and occ>=770 and occ<=975 then occ2=5;
        else if year<=2017 and occ>=980 and occ<=983 then occ2=5;
        else if year<=2017 then occ2=0;
if year>=2018 and occ>=10 and occ<=3550 then occ2=1;
        else if year>=2018 and occ>=3601 and occ<=4655 then occ2=2;
        else if year>=2018 and occ>=4700 and occ<=5940 then occ2=3;
        else if year>=2018 and occ>=6005 and occ<=7640 then occ2=4;
        else if year>=2018 and occ>=7700 and occ<=9830 then occ2=5;
        else if year>=2018 then occ2=0;
run;quit;

data dataF;
set data;
if female=0 then delete;
run; quit;


proc freq data =Data;
table sex race empstat nchild marst;
run;

Proc Means data =Data;
vars age inctot uhrswork;
run;

proc glm data= Data;
class race marst nchild citizen empstat occ2 ind2 year statefip silo ryear;
model inctot= female*policy female policy race age marst nchild citizen empstat occ2 ind2 uhrswork poverty
statefip year ryear silo Year*statefip/solution;
run; quit;

proc surveyreg data= Data;
class race marst nchild citizen empstat occ2 ind2 year statefip silo ryear;
model inctot= female*policy female policy race age marst nchild citizen empstat occ2 ind2 uhrswork poverty
statefip year ryear silo Year*statefip/solution;
run; quit;

proc glm data= Dataf;
class race marst nchild citizen empstat occ2 ind2 year statefip silo ryear;
model inctot= policy race age marst nchild citizen empstat occ2 ind2 uhrswork poverty statefip year ryear
silo/solution;
run; quit;

proc surveyreg data= Dataf;
class race marst nchild citizen empstat occ2 ind2 year statefip silo ryear;
model inctot= policy race age marst nchild citizen empstat occ2 ind2 uhrswork poverty statefip year ryear
silo/solution;
run; quit;

proc Surveyreg data= Data;
class race marst nchild citizen empstat occ2 ind2 year statefip silo ryear;
```

```
model inctot= female*policy female policy race age marst nchild citizen empstat occ2 ind2 uhrswork poverty
statefip year ryear silo Year*statefip/solution;
*where silo="CA";
*where silo="DC";
*where silo="NY";
*Where silo="NJ";
*Where silo = "RI";
where Silo="WA";
run; quit;

proc Surveyreg data= Dataf;
class race marst nchild citizen empstat occ2 ind2 year statefip silo ryear;
model inctot= policy race age marst nchild citizen empstat occ2 ind2 uhrswork poverty statefip year ryear silo
Year*statefip/solution;
*where silo="CA";
*where silo="DC";
*where silo="NY";
Where silo="NJ";
*Where silo = "RI";
*where Silo="WA";
run; quit;

proc glm data= Data;
class race marst nchild citizen empstat occ2 ind2 year statefip silo ryear;
model inctot= female*policy female policy race age marst nchild citizen empstat occ2 ind2 uhrswork poverty
statefip year ryear silo race*female age*female marst*female nchild*female citizen*female empstat*female
occ2*female ind2*female uhrswork*female poverty*female statefip*female year*female ryear*female silo*female
Year*statefip*female/solution;
run; quit;

proc surveyreg data= Data;
class race marst nchild citizen empstat occ2 ind2 year statefip silo ryear;
model inctot= female*policy female policy race age marst nchild citizen empstat occ2 ind2 uhrswork poverty
statefip year ryear silo race*female age*female marst*female nchild*female citizen*female empstat*female
occ2*female ind2*female uhrswork*female poverty*female statefip*female year*female ryear*female silo*female
Year*statefip*female/solution;
run; quit;

proc glm data= Data;
class race marst nchild citizen empstat occ2 ind2 year statefip silo ryear;
*model inctot= female*policy female policy race age marst nchild citizen empstat occ2 ind2 uhrswork poverty
statefip year ryear silo Year*statefip/solution;
Model inctot= female*policy female Policy;
where silo="CA";
run; quit;
```