Analyzing the American Gender Wage Gap: Why Do Different States Have Different Wage Gaps?

An honors thesis for the Department of Political Science

Reba Orloff

Tufts University, 2017

Abstract

A facet of the American gender wage gap is state level variation. Why are women paid 64 cents to the male dollar in Wyoming, but 89 cents to the male dollar in New York? Utilizing regression analysis, I explore my hypothesis that certain demographic and political variables have statistically significant relationships with state level gender wage gaps. My findings confirm my hypothesis, indicating that the Hispanic population, Black population, median age, urbanization, female share of the workforce, comparable worth legislation, and women in the state legislature are significant in their associations to state level wage gaps.

Acknowledgements

Thank you to Professor Berry for being an incredible advisor and mentor throughout the thesis process, as well as my college career. I am incredibly grateful to Professor Berry for his constant encouragement and belief in me. Thank you to Professor Masuoka for agreeing to be my second reader. Professor Masuoka's insightful feedback throughout the process benefited this project immensely, particularly the quantitative aspects. Thank you to Professor Eichenberg, whose seminar was extremely helpful in structuring and writing my thesis. Also thank you to Josh Quan for helping me with Stata. This thesis would not have been possible without the support of my friends, particularly Maddie, Amanda, Danielle, Ellie, Sarah, Emily, Stephen and Kirsten. Last, thank you to my family for always supporting me, especially my Mom: thank you for being the best and most inspiring role model I could ask for.

Table of Contents

Tables and Figures

Chapter 1: Introduction: Explaining American State Gender Wage Gap Variation

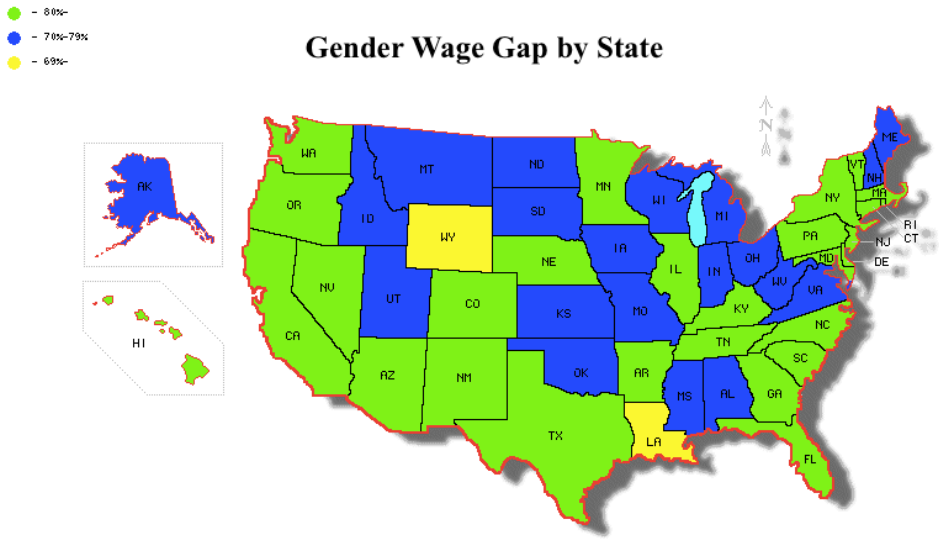
Although American gender equality has made progress, it is undeniable that gender discrimination remains prevalent. Inequality between the genders persists in many facets of American society. To name a few: in 2015 women made up 4.4 percent of Fortune 500 CEOs, and 19.4 percent of Congress ("List of Women," 2016; Rutgers University, 2015). What's more, the United States is the only industrialized nation without federal paid family leave (Livingston, 2016). Beyond issues of representation, parental leave, and countless other examples, there is one particularly pervasive aspect of gender discrimination in the United States: the gender wage gap, which has been said to "maintai[n] the lesser status of women in society" (Tharanou, 2012). In a discussion of gender equality, the wage gap is imperative. How can parity ever be reached if equal work does not constitute equal pay?

The gender wage gap, hereby defined as the difference in median dollars earned between all fulltime working men and working women, is a ramification of the historical and continuing gender discrimination in America, and is an issue commonly subject to political discourse. Although the national gender wage gap has narrowed since 1980, when it was 67 cents to the dollar, wages between the genders are far from equitable (Patten, 2015). In 2015, the White House reported that America's gender wage gap was such that women earned an average of 77 cents for every man's dollar, while Pew Research Center estimated the gap to be 84 cents per dollar (Patten). A complexity that these different estimates reflect is how to define the gap, which I will discuss in the literature review and methodology portions of the thesis. However, despite the many wage gap estimates and theories behind the discrepancy, the existence of a gender wage gap in the United States is undeniable.

Notably, the gender wage gap varies across the 50 states. In 2015, on opposite ends of the spectrum, the gender wage gap in Wyoming stood at 36 cents, and 11 cents in New York and Delaware ("Measuring the Wage," n.d.). Why is there a 25-cent range among the states in gender wage gaps? My thesis explores this question about state level gender pay gap variation.

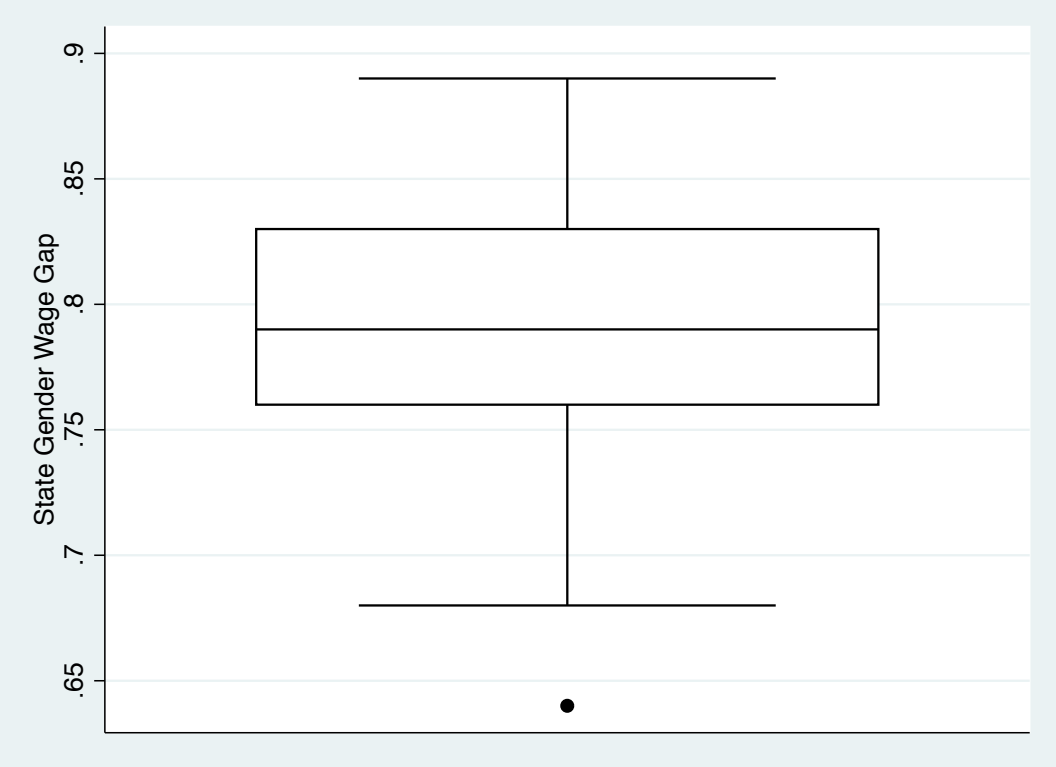
The spread of gender wage gaps across the states is depicted in Figure 1.1 below, created with the use of American Association University of Women (AAUW) data ("The Gender Pay Gap by State," 2015). Aside from New York and Delaware, Florida, North Carolina, and Rhode Island are among the states with relatively low wage differentials. Louisiana, West Virginia, Utah, and North Dakota are other states with particularly large gender pay gaps ("The Gender Pay Gap by State," 2015). Additionally, the dispersion of state wage differentials, including the 25th, 50th, and 75th percentiles, is visually depicted in Figure 1.2, a box plot.

Figure 1.1: Map of State Gender Wage Gaps



("The Gender Pay Gap by State," 2015)

Figure 1.2: State Gender Pay Gap Box Plot



(The Gender Pay Gap by State," 2015)

The topic matter of my thesis is crucial to today's world. Given the inequity that state and national gender wage gaps create, my thesis is firstly significant from a social justice perspective. It is important to comprehend why certain states are prone to larger gaps. Gaining knowledge of the factors most associated with state level wage inequality is a necessary step in working towards a normative model of equality.

Not only is state level gender pay gap variation important from a social justice perspective, but it is also significant to economic prosperity. It has been found that the gender wage gap and income per capita in the U.S. are inversely related. A 2016 study revealed, "a 50% increase in the gender wage gap leads to a 35% decrease in income per capita in the steady state" (Cavalcanti & Tavares). The gender pay gap is evidently a negative influence on the American

economy, and the economy can therefore benefit from further understanding of factors associated with the gap.

Last, my thesis relates to public policy. Because states have different amounts and types of equal pay policies, which I define as legislation aimed at preventing discrimination regarding the terms and conditions of employment between men and women, political variables are included in my analysis ("State Equal Pay," 2016). Policy is the most controllable means to narrow the gender pay gap, and through the policy component of my thesis, I hope contribute to the conversation about what equal pay policies are most effective.

I hypothesize that certain factors have statistically significant associations with the gender wage gap, and I approach my governing question with use of OLS regression. I include the following state level variables in my analysis, all of which I will discuss in detail in the thesis: industry makeup, racial makeup, existing equal pay legislation and parental leave policy, age composition, education level, partisan composition of state legislature, minimum wage, per capita income, whether there has been a recent vote on equal pay legislation, amount of women in the state legislature, urbanization, unionization, religiosity, fertility rate, and female participation in the workforce are inconsistent across states, and are associated with state level wage equality. Although I have hypotheses about the direction in which each of the included variables is related to the gender wage gap, each of which I will discuss in the thesis, this study is exploratory. I examine a range of plausible independent variables to measure their relationships with wage differentials.

It is notable that the gender wage gap is a heavily studied topic. On a national scale, experts have found and theorized a wide range of factors associated with the gap. Prior scholarship will be more thoroughly discussed in the literature review, Chapter 2, but scholars

have prominently emphasized gender segregation of industries, equal pay policy, and parental leave as factors most associated with the American gender pay gap (Blau & Kahn, 2000; Kim, 2013). I hope to add to the discussion with my analysis, particularly regarding the three commonly cited factors.

Significantly, to my knowledge, there is only one prior study of the gender wage gap using American states as the unit of analysis, and the gender wage gap as the dependent variable, and it seeks to explore a "state effect" of discrimination, rather than what factors are most significantly associated with the wage gap. Other studies of state level wage gaps are predominantly economic analyses that estimate state level wage gaps, or they are studies of a different scale, for example studies that utilize countries as the unit of analysis. My thesis will be the first such analysis of the gender wage gap, therefore contributing to scholarship of the gender wage gap with the new information gained. Overall, much can be learned about the American gender pay gap from analyzing factors associated with the gap across states.

In the following chapters, I explore the variation in gender wage gaps between the states. Chapter 2 discusses already existing analysis of the American and state level wage gaps, and what factors are most associated with discrepancies in pay between the genders. Chapter 3 focuses on my hypotheses, reasoning for the inclusion of the variables used in my study, and methodology for my regression analysis of the wage gap. Chapter 4 presents my models and findings. I end with Chapter 5, a summary of my findings, and discussion of study limitations and future implications.

Chapter 2: Literature Review of the American Gender Wage Gap

Introduction

It is undeniable that "scholarship in sociology and economics has long explored the gender wage gap," particularly within "wealthy countries," such as the U.S. (Misra and Murray-Close, 2014). My analysis of American state level gender wage gap variation builds on prior work. In this chapter, I review the existing literature on the American gender wage gap, and state level variation.

Fundamental Assumptions

Throughout the literature review, I discuss and analyze prior work on the governing research question of the thesis: why do different states have different gender wage gaps? This question involves assumptions based on prior studies, which must be reviewed.

The first assumption that my thesis is based upon is that the gender wage gap exists, and that it varies by states. This premise is substantiated by economists, Blau and Kahn, who confirm the existence of the gender pay gap in a study in which they refer to "data drawn from published government statistics on female-to-male earnings ratios" to establish that "women continue to earn less than men" (Blau & Kahn, 2007, p. 8-9). Another scholar, Hoffman, uses the same data to establish that the gender wage gap "varies substantially across U.S. states" (2013). The phenomenon of unequal pay is not new either: in 1963, the U.S. Department of Labor found that "lower pay rates for women doing the same work as men are not uncommon" (p.37). As assumed by Blau and Kahn and Hoffman, I argue that government recorded data substantiates the long held existence of the gender pay gap, and it's range across states.

The second fundamental assumption of my thesis, the definition of the gender wage gap, is far more nuanced. In a different work, Blau articulates the complexities involved in calculating

the gender wage gap with the statement that "there are many possible concepts that can measure pay" (2012, p. 190). Despite the difficulty associated with measuring the gap, in all of their work, Blau and Kahn employ the U.S. Census Bureau's definition: "median earnings of full-time year-round, workers" (U.S. Census Bureau, n.d.). As put by Blau, "the sample is restricted to full-time workers in order to focus on women and men whose commitment to employment is as similar as possible" (2012, p. 200). While it is significant that "women are more likely than men to work part-time" it is ultimately the best definition because "we do not have a similar long data series for hourly wages" that would be necessary to calculate the gap as "an hourly rate of pay" (Blau & Kahn, 2007, p. 8). The definition used by Blau and Kahn, and the U.S. Census Bureau, is evidently imperfect, but given the available data, it is the best definition of the gender wage gap. I assume the same definition in the thesis.

It is important to note another conception of the gender wage gap calculation, and why I disagree with it. According to Weichselbaumer and Winter-Ebmer, "the most common way to analyze [pay] discrimination based on gender" is the previously discussed method used by Blau and Kahn, but another procedure in economic analysis for calculating the pay gap was constructed in 1973 by Blinder-Oaxaca (2005, p. 481). The Blinder-Oaxaca method "allows that productive characteristics of men and women are rewarded differently," as opposed to "holding productivity constant" (Weichselbaumer & Winter-Ebmer, 2005, p. 481). In this context, productivity refers to characteristic gender differences, such as the fact that women tend to work fewer hours (Blau, 2012). I disagree with this logic of productivity, which holds that "not all observed differences in outcomes can be attributed to differences in opportunity," but rather can be attributed to personal preference ("Gender Equality and Development," 2012). Instead, as articulated by the 2012 World Bank "Gender Equality and Development" report, I posit that

"only by attempting to equalize outcomes can one break the vicious cycle of low aspirations and low opportunity." While certainly a matter of opinion, I argue that this alternate conception of the gender wage gap is flawed because productivity should be treated as a constant between men and women.

As I have discussed multiple studies calculation methods for the wage gap, the different model specifications are summarized in Figure 2.1 below. Figure 1.2 also provides the magnitude of the wage differential that is ultimately computed by each model. The degree to which the gender wage differential exists is a key element in the discussion of gender wage gaps.

Figure 2.1: Estimations of the American Gender Pay Gap

Author or Institution	Model Specification	American Gender Wage Differential	Year(s) Measured in the Differential
Hoffman	Earnings ratio for median annual earnings between full time male and female workers	.779	2008
Pew Research Center	Earnings ratio between men and women for median hourly wages of both part time and full time workers	.84	2015
Blau and Kahn	Earnings ratio for median annual earnings of full time male and female workers, unadjusted for covariates	.793	2010
Blau and Kahn	Earnings ratio for median annual earnings of full time male and female workers, adjusted for human capital related covariates	.821	2010
Blau and Kahn	Earnings ratio for	.916	2010

	median annual earnings of full time male and female workers, adjusted for all covariates		
Weichselbaumer and Winter-Ebmer	Earnings ratio for mean hourly wages between men and women, unadjusted for covariates	.74	1990-1999
Weichselbaumer and Winter-Ebmer	Earnings ratio for mean hourly wages between men and women, adjusted for covariates	.81	1990-1999

(Hoffman, 2013; Patten, 2015; Blau & Kahn, 2016; Weichselbaumer & Winter-Ebmer, 2005)

Framing of Studies

Prior to discussion of existing theories for state level gender wage gap variation, another necessary aspect of the literature to review is framings of related studies. Because to my knowledge there is no prior research that treats the same question with the same methodology, I must discuss what frameworks of gender wage gap studies have left out. Existing regression analyses on factors associated with the gender wage gap predominantly focus on meta-analysis, single-country analysis, or cross-country analysis. First, meta-analysis, a "tool to cumulate, review, and evaluate empirical research" has been conducted on an international and national scale (Weichselbaumer & Winter-Ebmer, 2005; Jarrell & Stanley, 2004). The scholars in both instances of meta-analysis compare wage gaps estimated by studies rather than the gap itself, which I contend is not the optimal framework for a study of state level variation. The most direct method to explore a question about the gender wage gap is to use the gap as the dependent variable. Second, many regression analyses of the gender wage gap focus on only the United States national pay differentials, or use countries as the unit of analysis in cross country analysis (Blau & Kahn, 2007). Although insightful into the American and global gender wage gaps, such

studies ignore the valuable information to be gained about the state and national wage gaps from analyzing state level variation. It is evident that existing framings of gender wage gap studies exclude state level regression analysis, which has potential to add to the existing literature.

Although it is a working paper, I must reference one study that attempts to fill the lack of regression analysis of state wage gaps. In a 2013 working paper, Hoffman explores why the gender earnings ratio is inconsistent across the U.S.. However, throughout the article, Hoffman focuses on the "state effect," which "summarizes the fact that a state's gender earnings ratio is higher or lower than expected, given" variables by states that have been shown to be associated with the pay gap (2013, p. 18). Hoffman uses related factors to the gender wage gap as controls to show that "the remaining effect must be related to differences across states in the value of traits or other unmeasured factors that differentially affect productivity by gender" (2013, p.18-19). Hoffman's study is fundamentally different from my thesis. Although a discrimination effect might provide insight into state level variation of the gender wage gap, I argue that it is not the best approach to an empirical study of state level gender pay gap variation. A discrimination effect is abstract, and a focus on concrete variables would logically provide the clearest results as to factors most associated with state level variation. It is evident that all prior framings, to my knowledge, are insufficient in exploring the governing question of the thesis, and my study must take a new approach.

It is lastly important to note another type of gender wage gap study framing that is insightful for the thesis: time series analysis. A 2005 study, by Weichselbaumer and Winter-Ebmer, provides an example of time series analysis of wage gap estimates by country. The study demonstrates that because the gender wage gap has narrowed, in a discussion of factors associated with the wage gap, it is logical to frame a question about the gap over time

(Weichselbaumer & Winter-Ebmer, 2005; Patten, 2015). I do not have the time capacity to include time analysis of the American gender wage gap, but it is important to recognize the value of such analysis.

Factors Associated with the Gender Wage Gap

As established, I approach my governing question with the subsequent question: what factors are associated with the gender wage gap, and what are my hypotheses for the direction of the relationships? Although the prior literature on this question, with the exception of Hoffman's study, focuses on the United States as a whole, I apply studies about the national pay gap to the formation of my model. The predominant theories in the literature about what factors are associated with the American gender pay gap are: gender segregation of industries, equal pay policy, and parental leave (Blau & Kahn, 2000; Kim, 2013). Although these theories are not the only theories about the gender wage gap, they are necessary to review in depth.

Gender Segregation of Industries

First, industry is a factor that has been found to have a relationship with gender pay differentials in America. While not a recent study, a 1995 work by Fields and Wolf highlights "differences between the patterns of inter industry wage differentials of men and women," a concept that is supported by a 2016 systematic review of the gender wage gap, as well as by Hoffman's working paper (p. 105; Bishu & Alkadry, 2016; 2013). Beyond different wage gaps by industry, industry has been posited to have a significant relationship with the gender wage gap because of gender segregation by industry, a trend in which "men are employed in higher paying occupations than women" (Misra & Murray-Close, 2015). This segregation purportedly begins to occur before careers even start. For example, Bobbitt-Zeher finds that the gender differences in "choice of a college major" patterns explain much of the gender wage gap in the entry-level

workforce (2007). Whether this segregation is a result of discriminatory tracking or preference-based differences, gender segregation of industries is important in a discussion of the gender wage gap. As industry and gender segregation are theorized by many scholars to be associated with the gender wage gap, and states have varying prominent industries, industry must be included as a variable in my model.

Another element of the existing literature that confronts gender segregation by industry is the conception of "human capital" as related to the gender wage gap (Misra & Murray-Close, 2014). This theory holds that gender differences in "education, experience and skills (also referred to as human capital)" have a significant relationship with the gender wage gap. One example of how human capital deficits between genders occurs is that "the traditional division of labor" means that "the longer hours women spend on housework decrease the effort they put into their market jobs compared to men" (Blau, 2012, p. 90). It is clear that the theory of human capital deficits having a positive association with the gender wage gap should be included in my study because different states have varying human capital levels. However, I argue that the manner in which prior works group human capital separately from industry gender segregation is flawed. These variables should not be grouped apart from industry segregation in analysis of the wage gap, because after all, the differences in human capital are inextricably linked to gender segregation of industries. If women are frequently responsible for a disproportionate amount of house duties, it is logical that as a group they might seek jobs with fewer hours. Thus, in my methodology and regression analysis, I will group human capital as a part of gender segregation by industry.

Equal Pay Policy

The second popular theory of variables associated with the gender wage gap is that equal pay policy is negatively related to the gap (Misra & Murray-Close, 2014). Equal pay policy as a remedy for the gender wage gap is widely discussed in the U.S. by scholars, politicians, and voters alike. In a 2014 floor speech U.S. Senator Udall from New Mexico supports the theory, stating, "I will keep working in the Senate to find a solution to pay inequality" ("Udall"). Scholars substantiate the idea of Udall's floor speech that legislation can be employed to successfully work toward equal pay. In a 2013 study, Kim finds that two main tactics of equal pay policy "increasing access to higher-paying jobs, and increasing pay in the jobs in which women already work." Kim asserts, the "research indicates that all of these policies improve women's wages and lower the gender wage gap" (2013). This theory is especially pertinent to a study of state level variation because states have different equal pay policies. I reason that state differences in policy are pertinent to my analysis because it has been said that in many states "the legal requirements for proving and winning compensation discrimination lawsuits are often unrealistically difficult" (Stanberry & Aven, 2013). It is logical that states with more comprehensive policy might have smaller wage gaps. Evidently, the theory of equal pay policy having a positive association to the gender pay gap is relevant to the thesis.

It must be mentioned, however, that a discussion of equal pay policy is complex due to the many iterations of anti pay discrimination legislation. Kim articulates the multifaceted nature of equal pay policies: "because the wage gap results from multiple causes, no single policy can end it, and multiple remedies are required" (2013). There are many types of equal pay policies instituted on a state level, all of which Kim argues have negative associations with the wage gap: "Title VII and affirmative action address the problem of women being employed in low-paying

jobs and being overlooked for higher-paid ones" and "comparable worth, unionization, and pay secrecy laws can allow women to increase their pay in the jobs in which they already work" (Kim, 2013). Thus, in my methodology and analysis, the existing literature stresses the need to be attentive to the details of policy types.

Parental Leave Policy

An additional area of policy that purportedly has a significant association with the gender wage gap is paid family leave. Pekkala Kerr contends that although "women's labor market histories display gaps in participation during years of family building," it has been found "that mandated parental leave improves women's economic position" (2015). In the existing literature, however, the consensus on parental leave policy is inconsistent. A 2014 study by an economist, Polacheck, concludes the opposite. Accordingly, Polacheck finds that "the gender wage gap in countries with greater paid family leave is larger" (2014, p. 8). Despite the disagreement in the literature about the direction of the relationship of parental leave policy with the gender pay gap, it is certain that because some states, such California and New Jersey, have "have specific mandates above and beyond the coverage of the FMLA" (federal Family and Medical Leave Act), parental leave policy must be included in my model (Pekkala Kerr, 2015, p. 1). Throughout the methodology and analysis portion of the thesis, I group equal pay policy and parental leave policy together, as they are all policy related variables.

Other Theories

In addition to the most prominently theorized factors associated with the American gender wage gap, the literature points to other variables that must be discussed. First, a 2010 study by Mason shows an association of racial identity with wages. Mason finds that there is a "persistent racial disparity among mature persons…in weekly wages during 1956-2006" (Mason,

2010). A 2013 study confirms the findings of a racial wage gap. The study states, "on average for all races, women are paid 77 cents for every dollar paid to men," however "African American women are paid only 70 percent of every dollar paid to men of all races and only 64 cents for every dollar paid to non-Hispanic White men" ("Study Finds," 2013). Evidently, the gender wage gap varies by race, demographics that are not constant across states, and therefore must be incorporated in this study of state level variation, a sentiment that is echoed by Hoffman's working paper (2013).

The minimum wage level has also been claimed by the literature to have a significant association with the American gender wage gap. Findings of 2011 study conducted by the U.S. Government Accountability Office (GAO), concludes that women "remain overrepresented among workers who earn low wages" ("United States Government," p. 8). Specifically, GAO "estimate[s] that women made up 49 percent of the overall workforce in 2010, but constituted 59 percent of the low wage workforce" ("United States Government," 2011, p. 8). As women evidently make up a disproportionate amount of the low-income workforce, lower minimum wage levels must logically be associated with higher gender wage differentials. This prior literature thus indicates the need to consider the minimum wage by state in my study.

Additionally, unionization has been found to have a significant relationship with the gender wage gap. Kim states that "the union wage premium is higher among women than among men in public sector jobs" which is important, as re-iterated by Jung and Cho, because increased rates of unionization are related to lower gender wage gaps (2013; 2014). As the U.S. department of labor reports that certain states have higher rates of unionization, for example New York is consistently the most unionized state, the literature implicates a need to include unionization in

my model ("Table 5." 2015). The need to include unionization in the model is also echoed by Blau and Kahn, who include unionization in their analysis of the gender wage gap (2016).

Lastly, there are theories about female participation in politics and better outcomes for women. These "critical mass" theories hold that "only as their numbers increase will women be able to work more effectively together to promote women-friendly policy change" (Childs and Krook, 2008, p.725). I therefore argue that consideration of the share of female participation in both the legislature and the workforce must be an element of my wage gap exploration.

I will last discuss the variables that Hoffman includes in his similar study that are not included in the thesis: marital status, number of children under 18, and self-employment. According to Hoffman, all of these variables are "factors and skill measures used regularly in the earnings literature" (2013, p. 9). Although I agree with Hoffman on the importance of these variables, given the high number of variables in my model, it not methodologically logical to use all of the variables that Hoffman does. The four variables that I have chosen to omit from my study can also be argued as confounding variables, such as number of children under 18 and marital status (Johnson et al., 2015; Pew Research Center, 2015).

However, it is notable that I have not included all variables in my model in this literature review, only the most prominent variables in the literature. Rationale for variables not discussed in depth in this chapter (age, education, partisan composition, income, existence of a recent equal pay policy vote, women in state legislature, urbanization, unionization, religiosity, fertility rate, and female share of the workforce) will be discussed in Chapter 3, the methodology portion of the thesis.

My Niche

It is clear that my thesis builds upon substantial scholarship on the topic of the American gender wage gap. As evident from my review of the literature on state level variation in the American gender pay gap, there is no prior study to my knowledge that analyzes the gender wage gap with the same approach. This thesis will therefore add to the scholarship of the gender wage gap by being the first regression analysis of state gender pay gaps as the dependent variable, testing for significance of factors that are reportedly associated with the gender pay gap, not including an abstract discrimination type effect.

The thesis is unique in the policy element of my analysis. Demographic variables associated with the gender wage gap are rarely discussed in the same studies as those that focus on equal pay policy. In the most similar study to my own to my knowledge, the scholar, Hoffman, states, "I have not…explored policies at the state level that may be related to the differences in the gender earnings gap and that is an obvious step for additional research" (Hoffman, 2013, p. 25). As pointed out by Hoffman, I argue that because state level demographic and policy factors have been found to have significant relationships with the gender wage gap, it is useful to explore them together. The thesis will therefore fill the lack of combined analysis of factors associated with the gender wage gap.

Overall, the thesis is unique in the framing of the governing question, which has only been approached by one other scholar, and in the combination of demographic type variables and policy variables. I hope that by filling in these gaps, knowledge will be gained of what factors are most significant in their association with state gender wage differentials, a necessary step in working toward equality.

Chapter 3: Hypotheses and Methods for Analyzing State Gender Wage Gaps

Introduction

I hypothesize that certain factors that vary across states are significantly associated with state gender wage gaps. Throughout this methodology-focused chapter, I specify the hypotheses of the models, outline how I operationalize each variable, and discuss the models I employ.

Hypotheses

As demonstrated in the wage gap literature, I hypothesize that industry makeup, racial makeup, existing equal pay legislation and parental leave policy, age composition, education level, partisan composition of state legislature, minimum wage, per capita income, whether there has been a recent vote on equal pay legislation, amount of women in the state legislature, urbanization, unionization, religiosity, fertility rate, and female participation in the workforce are inconsistent across states, and are associated with state level wage equality. This study is exploratory, and I therefore use a range of plausible independent variables, and test which ones are significantly associated with state gender wage gaps. In the following section of this chapter, each variable will be discussed. Further information on the data and data sources used in the study are located in the Appendix A.

Variables

State Gender Wage Gaps

My thesis uses the 50 American states as the unit of analysis, and the dependent variable is the gender pay differential in each state. The variable of state level wage gaps is defined as the "median annual…earnings ratio for full-time, year-round workers, by state and gender" ("The Gender Pay Gap," 2015). The dependent variable is expressed as the 2015 earnings ratio, ranging from .64 to .89 in Wyoming and New York respectively.

Industry Makeup

As discussed in Chapter 2, industry makeup in a state in states is an important consideration in analysis of state gender wage gap variation. However, given the wide array of industries, it is a complex variable to specify hypotheses for, and to quantify for the purposes of the thesis. For example, throughout their econometric analysis of the gender wage gap, Blau and Kahn use 14 different industry categories (2016). Given the constraints of time, degrees of freedom, and my statistical abilities, including similar industry variables in the study is not feasible. However, a simpler approach is demonstrate by Ryu, who in a 2010 study, posits that the industry variable most associated with the gender wage gap is the size of the public sector. According to Ryu, this is "because this sector creates lower-paid occupational positions for female employees and females may have limited authority or power in this employment setting" (2010, p.232). I therefore hypothesize that the percentage of the state working population employed by the public sector is negatively associated to wage equality among the genders. I also hold that the optimal quantification of industry makeup for this study is the percentage of the employed state population that works in the public sector, as of 2015.

Racial Makeup

Racial makeup of a state is an important variable because there is also a racial wage gap, where women that identify as African American and Hispanic experience the largest wage gaps compared to white men ("Study Finds, 2013). I hypothesize a negative relationship between racial diversity and equitable wages. Racial diversity is expressed in the study with three variables: the percentage in each state that identify white alone, the percentage in each state that identify non-Hispanic black, and the percentage in each state that identify as Hispanic, as of 2015. Although racial diversity is composed of far more than these three variables, the literature

suggests that these three are the strongest covariates with the wage gap, and given the low number of observations of 50 states, it is necessary to limit the independent variables (Hoffman, 2013).

Existing Policies, Equal Pay and Parental Leave

As prior work has highlighted the effectiveness of equal pay legislation, and some studies have suggest effectiveness of paid parental leave policy, I hypothesize a positive relationship between amount of existing legislation and wage equality among the genders. First, equal pay legislation refers to a wide spanning number of laws across the states, and thus must be simplified in the thesis. Given the predominant focus in the literature on comparable worth, defenses and remedies, and protections against pay discrimination, existing equal pay legislation is expressed as three dichotomous variables: whether or not each state had legislation falling into each category in 2014. Parental leave is expressed as a dichotomous variable of whether or not the state has paid family leave, as of the year 2015, although only three states do, California, New Jersey, and Rhode Island, and New York's will become effective in 2018 ("State Family and Medical," 2016).

Age Composition

Given that the literature indicates that younger generations experience lower gender wage gaps, I include age in the model (Patten, 2015). I hypothesize a negative association between age composition of the state population and wage equality, where in states with older populations women are paid less relative to men. In the thesis, the variable of age composition is expressed as the median age of the state population in 2015.

Education

Education levels are fundamental aspects of theories of human capital discrepancies and the gender wage gap. Therefore, I hypothesize that higher levels of in state education are positively associated with equitable wages among men and women. I operationalize education as one variable: percentage of state population 25 years old and over with a bachelor's degree or more, from 2015. While this variable does not capture all levels and elements of education, given the necessity to limit variables in the models, it represenst the variable that makes the largest difference in median national income (Blau and Kahn, 2007; "Median Weekly Earnings," 2014).

Partisan Composition of State Legislature

Given the party line tendency of legislatures on issues of equal pay policy, where the Democratic Party is in favor of legislation and the Republican Party opposes it, partisan composition of the state legislature must be considered. This variable is furthermore important to conceive on a cumulative basis, as the amount of time each party has been in power in a state is logically correlated to the amount of equal pay legislation that has been passed over time. I hypothesize that assuming that equal pay policies are effective in discouraging pay discrimination, Republican leaning partisan composition is negatively associated with wage equality. This variable is expressed an index score, covering the years 2006-2016, where for every two year period, a state receives one point for a Republican majority in either house, and zero points for Democratic majorities. In this index, the maximum index score is 30, and the minimum is zero.

Minimum Wage

As noted in the literature review, minimum wage level of states is relevant to the gender wage gap because women comprise a disproportionate amount of the workforce ("Gender Pay

Differences," 2011). I hypothesize that state mandated minimum wage levels are positively associated with wage equality, where a higher minimum wage is related to higher female wages relative to men. In the models, I use state minimum wages, expressed as dollar amount per one hour of work, from 2014.

Income

Although I hypothesize that state variation in per capita income is important to the gender wage gap, I include per capita income as an experimental variable. In the literature, there are conflicting findings on whether the gender wage gap is more or less substantial at higher income levels, which is further complicated by the exclusion of part time work in this study (Erosa et al., 2016). In the thesis, per capita income is expressed as the mean annual dollars earned per person in each state, in 5 year estimates ending with 2015.

Recent Vote?

I posit that whether or not the state legislature has recently voted on equal pay related legislation is related to the gender wage gap in the state. This is because existence or lack thereof of a recent vote is indicative of whether the gender wage gap is an issue of important in the state, regardless of whether the proposed legislation passed. I hypothesize recent votes are positivelu related to higher wages to women relative to men. In the thesis, this variable is as a dichotomous variable that represents if there has been a vote in the last two complete state legislative session, 2014 and 2015, where 0 indicates no vote, and 1 indicates a vote.

Women in the State Legislature

Scholars have found that women are more likely than men to propose, sponsor, and vote for equal pay legislation (Wolbrecht, 2000). I hypothesize that the larger the amount of women in the state legislature is, the more equal pay legislation the state will have, and the more

equitable wages will be. The expression of this variable in the thesis is the percentage of total seats in the state legislature that were held by women in 2014.

Urbanization

As articulated by Hoffman, it has been found that "family location decisions are made with priority given to male employment prospects" (2013). Thus, I hypothesize a positive association between urbanization and wage equality, because if more of the state population is urban dwelling, limited employment prospects that disproportionately impact women will be less prevalent. I express urbanization as the percentage of the state that lived in an urban environment in 2010. Urban dwelling is defined using the Census Bureau's classifications ("Urban and Rural," 2010).

Unionization

As women in unions have higher wage premiums than men in the public sector, I hypothesize a positive association between state unionization rates and wage equality among the genders (Kim, 2013). In the dataset, the unionization variable is represented by the percentage of the state's employed population that is represented by unions, from 2015.

Religiosity

Given that high religiosity has been found to be positively associated with gender inequality, I posit that the religious tendencies in a state are relevant to the gender wage gap, a form of gender inequality (Seguino, 2011). I hypothesize that high religiosity is negatively associated with equitable wages. In the models, I express religiosity with the 2016 measures of the Pew Research Center, which estimates religiosity with an overall index number comprised of religious importance, frequency of prayer, worship attendance, and belief in god, expressed as percentages that range from .77 to .33 among the states (Lipka, 2016).

Fertility Rate

Childcare has been found to have a positive relationship with the gender wage gap, and I argue that the number of women children women have, ranging from 1.56 to 2.33 in the states, is related to the gender wage gap. I hypothesize that higher state fertility rates are negatively associated with equitable pay among the genders. In the models, fertility rate is represented by the average number of children women have in their lifetime within a state, as of 2014.

Female Share of Workforce

As discussed in Chapter 2, there are theories of a connection between greater female participation in workplaces, particularly in leadership positions, and better outcomes for women (Wolfe and Fields, 1995). I hypothesize that higher shares female-identifying members of the workforce is positively associated with wage equality. In the models, female share of the workforce is expressed as the percentage of the full time workforce that is female.

Model Estimation and Data Transformation

I use the computer software, Stata, to analyze my data and generate my findings. It is important to note that the type of model employed in the study is ordinary least squares regression (OLS). OLS regression is the best model for the study because all of the independent variables, some with changes to the data, fit the requirement of OLS regression for normal distribution, and have seemingly linear relationships with the dependent variable (Kellstedt & Whitten, 2009). Accordingly, the following changes were made to the data in order to fit the assumption of normality: the percentage of the full time employed population in the public sector and fertility are inversed, the percentage of the population identifying as Black alone, the percentage of the population identifying as Hispanic alone, education, urbanization, and per capita income are all expressed as natural logarithms, median age and female share of the

workforce are squared, the percentage of the population identifying as white is cubed, because the partisanship index is unable to be transformed to a normal variable, it is expressed as a dichotomous variable where a 1 represents above median scoring on the partisanship index, and 0 represents a below median score. Lastly, the minimum wage variable is omitted due to the inability of the variable to transform to normal distribution. OLS is also the optimal regression model given the use of it in much of the prior work gender wage gap work, for using the same regression model make studies most comparable (Hoffman, 2013).

Another notable aspect of the model is that although this is a cross sectional study, the data is drawn in from different years, ranging from 2010 to 2016. I argue, however, that the discrepancies in the data are admissible because the data is drawn from the most recent plausible sources, and the range of years is relatively low. Further explanation and discussion of year variation in the data sources can be found in Appendix A.

Finally, given the large quantity of variables in the study, I will use multiple models to explore state level variation in gender pay differentials. In the Chapter 4, I specify and explain the composition of the initial model, and subsequent different models.

Summary of Variables and Hypotheses

Given the quantity of variables discussed in this chapter, Figure 3.1 below summarizes the hypotheses and measures of variables outlined in this section, and notes any transformations made to the data to fit the OLS requirement of normal distribution.

Figure 3.1: Summary of Hypotheses and Measures

Variable	Hypothesis	Measure
Industry Makeup	Public sector employment is positively associated with wage equality	Percentage of state full time employed population in the public sector, inversed for normal distribution

Racial makeup	Racial diversity is negatively associated with wage equality	1) Percentage of state population that identify as non-Hispanic white, cubed for normal distribution. 2) Percentage of state population that identify as non-Hispanic black, expressed as the natural logarithm for normal distribution. 3) Percentage of state population that identify as Hispanic, expressed as the natural logarithm for normal distribution
Existing equal pay and parental leave legislation	Legislation is positively associated with wage equality	1) Existing comparable worth legislation, 0-no, 1-yes 2) Existing defenses and remedies legislation, 0-no, 1-yes 3) Existing protections legislation, 0-no, 1-yes 4) Existing paid family leave 0-no, 1-yes
Age composition	Lower ages are positively associated with wage equality	Median age of state population. Squared for normal distribution
Partisan composition of legislature	The Republican Party is negatively associated with wage equality	10 year index, where every two years, one point is assigned for a Republican majority in either chamber of the state legislature, expressed as 0-below median index score, 1-above median index score for normal distribution
Education	Higher education levels are positively associated with wage equality	Percentage of state population 25 and up with bachelor's degree or more, expressed as the natural logarithm for normal distribution.
Minimum wage	Minimum wage requirements are positively associated with wage equality	Legal minimum hourly wage requirement, discarded due to an inability to transform to normal distribution
Income	Experimental	Per capita income, expressed as the natural logarithm for

		normal distribution
Recent vote	A recent vote is positively associated with wage equality	0-there has not been a vote, 1-there has been a vote
Women in legislature	Women in the state legislature are positively associated with wage equality	Percentage of seats in state legislature held by women
Urbanization	Urbanization rates are positively associated with wage equality	Percentage of population that lives in urban environments. Expressed as the natural logarithm for normal distribution.
Unionization	Unionization rates are positively associated with wage equality	Percentage of workforce that is represented by unions
Religiosity	Religiosity is negatively associated with wage equality	Religiosity index score (0%-100%)
Fertility rate	Fertility rate is negatively associated with wage equality	Average number of kids per female lifetime. Inversed for normal distribution.
Female share of workforce	Female participation in the workforce is positively associated with wage equality	Percentage of the full-time workforce 16 and over that is female

Chapter 4: Findings and Analysis: Why Do Different States Have Different Wage Gaps?

Introduction

In this chapter, I utilize the dataset to test my hypotheses about state level gender wage gaps and the included demographic and political variables. I rely on both correlations and regression analysis.[1]

Correlations

The first notable finding is the correlations, or covariances, of the independent variables with state gender pay gaps. The correlations, both in direction and magnitude, have implications for the relationships of the variables with the wage gap. Figure 5.1 below displays each independent variable's correlation with the gender pay gap. In instances where I transformed variables for normal distribution, both the original and transformed correlations are included. Although these findings measure the degree to which two random variables vary together without controlling for external factors, the findings are noteworthy (Kellstedt and Whitten, 2009).

Figure 5.1: Correlation of Independent Variables with the Wage Gap

Variable	**Correlation to Wage Gap**
1) Industry makeup 2) Industry makeup (inversed)	1) -0.2792 2) 0.2599
1) Non-Hispanic white population 2) Non-Hispanic white population (cubed) 3) Non-Hispanic black population 4) Non-Hispanic black population (logged) 5) Hispanic population 6) Hispanic population (logged)	1)-.3802 2)-0.3498 3) 0.1574 4) 0.309 5) 0.4209 6) 0.4440
1) Existing comparable worth legislation, 0-no, 1-yes	1)-0.1873

[1] Note that because the dataset expresses the wage gap as the percentage of female to male wages, throughout my analysis a positive association to the wage gap indicates lower wages for women relative to men. For ease of comprehension I largely refer to associations with "wage equality," where a positive association represents higher wages for women.

2) Existing defenses and remedies legislation, 0-no, 1-yes 3) Existing protections legislation, 0-no, 1-yes 4) Existing paid family leave 0-no, 1-yes	2) 0.0407 3) 0.0118 4) 0.2692
1) Median age of state population 2) Median age of state population (squared)	1) 0.2463 2) 0.2333
1) Partisanship composition index 2) Partisanship composition index (dichotomous: 0-below median, 1-above median)	1)-0.4318 2) -0.3995
1) Education 2) Education (logged)	1) 0.3937 2) 0.3987
Minimum wage	0.3751
1) Income 2) Income (logged)	1) 0.2877 2) 0.2901
Recent vote?	0.0660
Women in legislature	0.4715
1) Urbanization 2) Urbanization (logged)	1) 0.4308 2) 0.3820
Unionization	0.2880
Religiosity	-0.0681
1) Fertility rate 2) Fertility rate (inversed)	1)-0.4584 2) 0.4402
1) Female share of workforce 2) Female share of workforce (squared)	1) 0.5209 2) 0.5219

The direction of correlations with the wage gap, expressed in Figure 3.1, are consistent with my hypotheses, except for the industry makeup, racial diversity, comparable worth legislation, and median age. Where the relationships between independent variables and the wage gap that counter my hypotheses are found to be significant, I present potential explanations in the "Results" portion of this chapter. For the one variable included as experimental, income, the correlation reflects a positive association between per capita income and equitable wages, suggesting that states with higher per capita income have smaller pay differentials. In all instances of transformed variables, the direction of the correlations are consistent for both variables, aside from inversed variables, which predictably flip direction. Interestingly, the two variables with the highest covariance with the pay differential, both positively correlated with wage equality, are the percent of the female share of the workforce, and women in the state legislature (.5209 and .4715 respectively). As the magnitude of the correlations suggest, these

two variables prove to be significant in my regression analysis. In the "Results" portion of this chapter, I explain the significance of these two variables, and in the "Conclusion" I consider the potential statistical issue of reverse causality as a result of the strong correlations.

Model Composition

Given the quantity of independent variables and relatively low number of observations, the thesis required multiple models, and the elimination of certain variables. The statistical error of multi-collinearity between the independent variables was another critical consideration in model composition (Kellstedt and Whitten, 2009). Adhering to a correlation magnitude cut off of .6 or greater, I determine the following variables to be collinear with one another, both before and after transformation where applicable.[2]

1. Fertility rate and median age
2. Religiosity and per capita income
3. Religiosity and women in state legislator
4. Education and per capita income
5. Women in state legislature and education
6. Religiosity and education
7. Urbanization and Hispanic population
8. Female share of workforce and Black population

To avoid statistical error, at least one of the variables in each of the above relationships is omitted from the models, except Model II, which uses collinear variables for exploratory reasons discussed in the "Results" section of this chapter. I explain the decisions made in each model specification in detail below, but note that fertility rate and religiosity are left out of all models. Due to flaws in year of collection and objective difficulty of the variable to quantify of the

[2] The only relationship that is not found to be above 0.6 with both the original and transformed versions of the variables is the correlation between the Black population and the female share of the workforce. However, the correlation prior to transformation of both variables is .5989, and above .6097 after transformation. I therefore categorize the relationship as collinear.

fertility and religiosity variables respectively, I favor the collinear variables. In order use the female share of the legislature where political variables are included, and maintain consistency in demographic variables, throughout all models I use only income, opposed to education. Finally, the minimum wage is also omitted from all model compositions because it is abnormally distributed and incapable of normal transformation. Further information on collinearity, and the statistical limitations it creates for the study, are located in Appendix B and Chapter 5 respectively.

Below are the specified models. Figure 5.2 indicates the p-values for the variables included in each model.

Model I: y_i (wage gap)= â + B(inverse industry makeup)+ B(square age composition)+ B(log income)+ B(unionization)+ B(log Black)+ B(log Hispanic)

Model II: y_i (wage gap)= â + B(inverse industry makeup)+ B(square age composition)+ B(log income)+ B(unionization)+ B(log Black)+ B(log Hispanic))+B(log urbanization)+ B(square female share of workforce)

Model III: y_i (wage gap)= â + B(inverse industry makeup)+ B(square age composition)+ B(log income)+ B(unionization)+ B(log Black)+ B(log Hispanic)+ B(paid family leave) + B(comparable worth) + B(defenses and remedies) + B(protections) + B(recent vote) + B(partisanship composition) + B(women in legislature)å

Model IV: y_i (wage gap)= â + B(inverse industry makeup)+ B(square age composition)+ B(log income)+ B(unionization)+B(log urbanization)+ B(square female share of workforce)

Model V: y_i (wage gap)= â + B(inverse industry makeup)+ B(square age composition)+ B(log income)+ B(unionization)+ B(log urbanization)+ B(square female share of workforce)+ B(paid family leave) + B(comparable worth) + B(defenses and remedies) + B(protections) + B(recent vote) + B(partisanship composition) + B(women in legislature)

Model VI: y_i (wage gap)= â + B(paid family leave) + B(comparable worth) + B(defenses and remedies) + B(protections) + B(recent vote) + B(partisanship composition) + B(women in legislature)

Figure 5.2: Summary of Models

	Model I	Model II	Model III	Model IV	Model V	Model VI
Unionization	.325	.196	.551	.351	.850	__
Industry makeup (inversed)	.900	.159	.926	.605	.796	__
Income (logged)	.683	.664	.634	.951	.321	__
Median age (squared)	.000**	.929	.050**	.603	.856	__
Black population (logged)	.066*	.156	.002**	__	__	__
Hispanic population (logged)	.000**	.000**	.034**	__	__	__
Urbanization (logged)	__	.213	__	.012**	.008**	__
Female share of workforce (squared)	__	.000**	__	.000**	.000**	__
Family leave	__	__	.761	__	.163	.238
Comparable worth	__	__	.099*	__	.035**	.001**
Defenses and remedies	__	__	.723	__	.384	.542
Protections	__	__	.632	__	.991	.478
Recent vote	__	__	.555	__	.533	.453
Partisan composition (dichotomous)	__	__	.530	__	.522	.066*
Women in legislature	__	__	.018**	__	.006**	.017**
R^2	.4881	.6841	.6272	.5040	.6843	.3999
N	50	50	49	50	49	49

*p<.1 **p<.05

Results

Model I: y_i (wage gap)= â + B(inverse industry makeup)+ B(square age composition)+ B(log income)+ B(unionization)+ B(log Black)+ B(log Hispanic)

In Model I, I analyze only demographic variables. Hypothesizing that racial makeup must be included in my regression due to the magnitude of gender wage gaps when intersected with racial wage gaps, I include the racial variables and exclude the collinear variables of urbanization and female share of workforce. Model I is significant at the 5 percent level, and has an R squared value of .4881.[3]

Figure 5.3: Summary Model I Regression Results

Variable	**Coefficient**
Unionization	.0014825
Industry makeup (inversed)	.0011743
Income (logged)	.0228854
Median age (squared)	.0001239**
Black population (logged)	.0125197*
Hispanic population (logged)	.00338046**

*p<.1 **p<.05

Model I's output first indicates that the natural logarithms of both the Black and Hispanic populations are positively associated with equitable wages. The natural logarithm of the Black population is significant at the ten percent level, and the natural logarithm of the Hispanic population is significant at the five percent level. A one unit increase in the natural logarithm of the state Black population is associated with a .0125 increase in the ratio of female to male

[3] After detecting a degree of heteroskedasticity in the distribution of the residuals in every model discussed in this chapter, I utilized the Stata command, "robust," to correct for the statistical issue ("Robust Regression," n.d.). All models included are therefore robust regressions, which notably do not include adjusted R squared values. All R squared values discussed are unadjusted.

wages, and a one unit increase in the natural logarithm of the Hispanic population is associated with a .0338 increase in the ratio of female to male wages.

This finding counters my hypothesis that racial diversity corresponds to larger state level wage gaps, as on a national scale Black and Hispanic women are reportedly paid at respective rates of 70 and 64 percent compared to white men ("Study Finds," 2013). I postulate that this finding may be because states with higher racial diversity are less prone to discrimination of all sorts. Another potential explanation for the significance of the Black and Hispanic population variables is that it is spurious. Perhaps the relationship of the racial variables with omitted collinear variables drives the result, or issues within the data: that part time workers are unaccounted for, or that the racial demographics are representative of the entire state population, opposed to the full time workforce. Although given the constraints of the study I cannot adjust for potential data flaws, it is possible to explore collinear relationships of the racial variables, which I do in subsequent models.

Model I also indicates that age variable is significant and positively associated with wage equality. It is significant at the five percent level, and a one-year increase in the age variable is associated with a .0001 increase in the ratio of female to male wages. The magnitude of the relationship is evidently miniscule, but this finding is again counter to my hypothesis of a negative relationship, as studies demonstrate that younger women are paid more equally to their male counterparts, and that women retire at younger ages than men ("Employed Persons," 2017). I hypothesize that states with higher median ages are home to industries that pay more equitably. Another possible explanation is that state median age does not accurately cover the age distribution of the workforce, and states with higher median ages might actually have more young members of the full time workforce. Alternatively, as with the racial findings, there might

be spurious drivers of this finding, such as multi-collinearity, or a complication within the data. I use Model II to further explore the findings of Model I.

Model II: y_i (wage gap)= â + B(inverse industry makeup)+ B(square age composition)+ B(log income)+ B(unionization)+ B(log Black)+ B(log Hispanic))+B(log urbanization)+ B(square female share of workforce)

Model II represents the same demographic model as Model I, with the variables of urbanization and female share of the workforce included. Although there is multi-collinearity with the racial makeup and two added variables, I include all four in order to further explore the difficult to explain racial diversity findings of Model I. This model is significant at the five percent level, and has a notably high R squared value, .6841. A high r-squared value can be indicative of spurious results, and is particularly relevant to this study given that there are only 50 cases, and therefore relatively low degrees of freedom. However, as described in Footnote 3, the fact that the R squared value is unadjusted implies that this value would be lower if an adjusted value was available. Overall, the high R squared does not render this model insignificant to the study, but it is an important limitation to bear in mind.

Figure 5.4: Summary Model II Regression Results

Variable	**Coefficient**
Unionization	.001739
Industry makeup (inversed)	.0131467
Income (logged)	.0204758
Median age (squared)	2.86e-06
Black population (logged)	-.0066884
Hispanic population (logged)	.045824**
Urbanization (logged)	-. 0689323
Female share of workforce (squared)	.0002609**

*p<.1 **p<.05

The first finding of Model II is that the natural logarithm of the Hispanic population remains significant, at the ten percent level, such that a one unit increase in the natural logarithm of the Hispanic population is associated with a .0458 increase in female wages relative to male wages. However, the natural logarithm of the Black population does not remain significant in this model. The second finding of Model II is that the squared female share of the workforce is significant at the ten percent level. A one-unit increase in the squared female share of the workforce is related to .0003 higher wages for women relative to men. Notably, the urbanization variable is insignificant in Model II.

The finding of the Hispanic population's significance reinforces the strong association between lower wage gaps, and larger Hispanic populations. As within Model I, I hypothesize that states with larger Hispanic populations might have greater concentrations of jobs industries with specifically high wage equality, or that greater racial diversity might be indicative of less discrimination of all forms in a state.

As hypothesized in Chapter 3, the finding that the female share of the workforce is positively associated with lower pay differentials between men and women is consistent with multiple theories of gender discrimination. It is theorized that greater female participation in the workforce normalizes female employment, and as women reach greater mass, they will have more ability to work toward equal pay as a group, particularly when women increasingly hold positions of power (Browne, 2014).

Overall, despite admitted statistical flaws, Model II demonstrates the strength of the relationships between state gender wage gaps and both the Hispanic population and the female share of the workforce. Because one of the racial variables and one of the collinear pairs with the racial variables are significant when all four are included in the model, I argue that neither pair is

necessarily preferable for omission from the models. I therefore test political variables with two different demographic models.

Model III: y_i (wage gap)= â + B(inverse industry makeup)+ B(square age composition)+ B(log income)+ B(unionization)+ B(log Black)+ B(log Hispanic)+ B(paid family leave) + B(comparable worth) + B(defenses and remedies) + B(protections) + B(recent vote) + B(partisanship composition) + B(women in legislature)å

In Model III, I extend Model I to include political variables. This model is significant at the five percent level, and has an R-squared value of .6272, which as with Model II, is notably high. A feature of Model III is that there are 49 observations, opposed to 50, because Nebraska does not have a bicameral legislature, and therefore cannot be scored in the partisanship index ("States-Ballotpedia," 2016).

Figure 5.7: Summary Model III Regression Results

Variable	**Coefficient**
Unionization	.0010534
Industry makeup (inversed)	-.0006261
Income (logged)	-.0301868
Median age (squared)	.000082**
Black population (logged)	.0208611**
Hispanic population (logged)	.0196542*
Family leave	. 0080122
Comparable worth	-.0262739*
Defenses and remedies	-.004475
Protections	.006086
Recent vote	-.00721
Partisan composition (dichotomous)	-.0095216
Women in legislature	. 0030064 **

*p<.1 **p<.05

The first demographic finding of Model III is that the squared median age is significant at the five percent level, such that a one-unit increase in the squared median age corresponds to .000082 higher wages for women as compared to men, a markedly minimal regression

coefficient. Second, the natural logarithm of the Black population is significant at the five percent level, where a one-unit increase in the logged Black population is related to .0209 more equitable wages. The natural logarithm of the Hispanic population is also significant at the ten percent level. A one-unit increase in the natural logarithm of the Hispanic population is associated with .0197 more equal wages between the genders. I hypothesize the same explanations for the significance of all three of these variables discussed in Model I, and argue that the continued significance of these variables when political variables are accounted for indicates the significance of these demographic variables to the gender wage gap.

The first political finding of Model III is that comparable worth legislation is significant at the ten percent level. Comparable worth legislation is negatively associated with wage equality such that states with legislation enacted correspond to a .0263 wider wage gap. This finding is counter to my hypothesis that all equal pay legislation is positively related to wage equality. However, I argue that the five states with comparable worth legislation (Illinois, Iowa, Minnesota, Montana, and West Virginia) are predominantly liberal states with below median gender pay gaps.[4] I theorize that these states are more likely to pass equal pay legislation as the correlations of the legislative variables with the partisanship variable indicate, and because these states are more in need of equal pay legislation due to higher wage inequality. Another element of this finding is that comparable worth legislation itself is not effective as legislation. Comparable worth was a popular movement in the 1980s that aimed to "require an employer to eliminate a negative partial correlation between wages and femaleness of occupation," and has since been deemed relatively difficult to enforce (Johnson and Solon, 1986, p.1117). According

[4] All states with comparable worth legislation have below median partisanship index scores with the exception of Montana, and all five states have wage gaps below the median, with the exception of Illinois that has the median wage gap, and Minnesota, which is above the median.

to scholars, Johnson and Solon, comparable worth "eliminate[s] only a small fraction of the gap between average male and female wages" (Johnson and Solon, 1986, p.1124). Overall, I assert that comparable worth is significant and negative for wage equality not because this legislation creates inequitable wages, but because these states are more prone to pass equal pay legislation than others given their liberal legislatures and high wage inequality, and comparable worth is ultimately not an effective tactic for enforcing equal pay.

The last finding of Model III is a positive association between women in the legislature and wage equality, significant at the five percent level. A one percent increase in the percent of the state legislature that is female is related to a .0030 increase in female wages relative to males. This finding works to confirm the "critical mass" theory of descriptive representation of women discussed in the literature review, which holds that increased presence of women in legislatures corresponds to better outcomes for women (Childs and Krook, 2008). Additionally, although the dichotomous legislation variables are found to be mostly insignificant, I do not account for the actual content of the legislation beyond the categories that legislation falls into. Perhaps larger amounts of female legislators create legislation content, both equal pay and otherwise, that leads to better outcomes for women.

Model III communicates the significance of racial diversity, age, comparable worth, and female presence in the state legislature to the gender wage gap.

Model IV: y_i (wage gap)= â + B(inverse industry makeup)+ B(square age composition)+ B(log income)+ B(unionization)+B(log urbanization)+ B(square female share of workforce)

As previously stated, I use Model IV to represent the same demographic regression as Model I, but with the noted swap of collinear variables, the Hispanic and Black populations with urbanization and female share of workforce respectively. This model is significant at the five percent level, and has an R squared value of .5040.

Figure 5.6: Summary Model IV Regression Results

Variable	**Coefficient**
Unionization	.0012597
Industry makeup (inversed)	.0044636
Income (logged)	.00312
Median age (squared)	.0000198
Urbanization (logged)	.0961391**
Female share of workforce (squared)	.0001967**

*p<.1 **p<.05

Model IV indicates that the natural logarithm of urbanization and the squared female share of the workforce are significant at the five percent level, and positively associated with lower wage differences between men and women. A one-unit increase in the natural logarithm of urbanization is associated with a .0961 increase in the female to male wage ratio, and a one-unit increase in the squared female share of the workforce corresponds to a .0002 increase.

As hypothesized in Chapter 3, it is logical that wage inequality between men and women is less pervasive in urban locations. Within urban areas, women have greater access to higher paying jobs for which they qualify, and are not as impacted by deprioritization in familial decision to relocate for employment as their rural counterparts (Hoffman, 2013).

I hold the same explanation for the finding of greater wage equality with more women in the workforce as within Model II. Furthermore, that greater female participation in the workforce normalizes female employment, and as women reach greater mass, they will have more ability to work toward equal pay as a group, particularly when women increasingly hold positions of power (Browne, 2014).

Overall, Model IV is evidence of the importance of the relationships between wage equality and both urbanization and female participation in the workforce.

Model V: y_i (wage gap)= â + B(inverse industry makeup)+ B(square age composition)+ B(log income)+ B(unionization)+ B(log urbanization)+ B(square female share of workforce)+ B(paid family leave) + B(comparable worth) + B(defenses and remedies) + B(protections) + B(recent vote) + B(partisanship composition) + B(women in legislature)

In Model V, I extend Model IV to include political variables. Model V is significant at the five percent level, and has an R squared value of .6843, which similarly to Model II and Model III, is notably high. As with all other models including political variables, Model V has 49 observations.

Figure 5.7: Summary Model V Regression Results

Variable	**Coefficient**
Unionization	.0003271
Industry makeup (inversed)	.0018005
Income (logged)	-.0528953
Median age (squared)	8.55e-06
Urbanization (logged)	. 0816669**
Female share of workforce (squared)	.0001988**
Family leave	. 0355729
Comparable worth	-.0246944 **
Defenses and remedies	-.0107075
Protections	-.0003742
Recent vote	-.0054412
Partisan composition (dichotomous)	-.0105527
Women in legislature	. 0030064 **

*p<.1 **p.<.05

The demographic results of Model V reflect the results of Model IV. Both the natural logarithm of urbanization and the squared female share of the workforce are positively associated with equitable wages, and are both significant at the five percent level. A one-unit increase in the natural logarithm of urbanization is associated with a .0817 increase in the female

to male wage ratio, and a one unit-increase in the squared female share of the workforce is associated with a .0002 increase in the female to male wage ratio. I assert the same explanations for the significance of these variables as in Model IV, and that the enduring significance of the variables with political controls indicates the importance of urbanization and female share of the workforce in state gender wage gap differentiation.

The political finding of Model IV reflect that of Model III: comparable worth legislation is significant at the five percent level, where it is negatively associated with wage equality such that states with legislation enacted correspond to a .0247 wider wage gap. Women in the state legislature is again significant at the five percent level. A one percent increase in the percent of the state legislature that is female is related to a .0030 increase in female wages relative to males. I posit the same explanations for the significance of these specific political variables as within Model III.

Model V enforces the finding that higher participation of women, in the workforce and in the state legislature, relates to less wage inequality.

Model VI: y_i (wage gap)= â + B(paid family leave) + B(comparable worth) + B(defenses and remedies) + B(protections) + B(recent vote) + B(partisanship composition) + B(women in legislature)

Last, to explore the relationships of the political variables with the wage gap, I utilize Model VI to explore only political variables. Model VI is significant at the five percent level, and has an R squared value of .3999. While this model is admittedly flawed in its omission of all demographic controls, the findings help to solidify the findings of Model VI. As with Model III and Model V, Model VI has 49 observations.

Figure 5.8: Summary Model VI Regression Results

Variable	**Coefficient**
Family leave	.0257804
Comparable worth	-.0515857 **

Defenses and remedies	-.0090866
Protections	.0098656
Recent vote	.0117506
Partisan composition (dichotomous)	-.0306652*
Women in legislature	.0030893**

*p<.1 **p<.05

Comparable worth and women in the legislature are again significantly related to wage equality between the genders, both at the five percent level. Existing comparable worth legislation is associated with a .0516 lower female to male wage ratio. A one percent increase in women in the legislature is associated with a .0031 greater female to male wage ratio. I assert the same explanation for the significance of these two variables as within Model III and Model V.

Last, above median partisanship index scores for Republican dominance of the state legislature is significant at the ten percent level, such that states with above median scores on the partisanship index created for the study are associated with .0307 lower female to male wage ratios. This finding reflects the negative correlation between the partisanship index and all other political variables included, aside from defenses and remedies legislation.[5] Assuming that equal pay legislation, amount of women in the state legislature, and recent votes on equal pay legislation are effective to some degree in lowering the pay gap, it is not surprising that Republican dominated states are associated with larger pay differentials.

Although Model VI is admittedly speculative due to lack of controls, this model demonstrates consistent significance of the political variables of comparable worth legislation

[5] Above median scores on the partisanship index are positively correlated with defenses and remedies legislation to a degree of .1. While this correlation is positive, it is arguably low in magnitude.

and women in the state legislature across all three models that include political measures. These two political variables are important considerations in discussion of the wage gap.

Conclusion

Overall, my models and findings demonstrate the significance of Hispanic population, Black population, median age, urbanization, female share of workforce, comparable worth legislation, and female share of the legislature to state level gender wage gaps. Hispanic population, female share of the workforce, Comparable Worth legislation, and women in the state legislature are especially important, as these variables endure as significant across all models in which they are included. While this study was experimental with regard to what variables might be significant, my hypothesis that some variables are significant in their associations with state level gender wage gaps is confirmed.

A key pattern of the findings is the aspect of female participation, both in the workforce and the state legislature. My analysis supports the theory that female inclusion in the labor markets and political office leads to better outcomes for women, at least in terms in wage equality. However, given the strong correlations with these two variables and state gender pay gaps (.5209 for female share of the workforce, and .4715 for women in the state legislature), the potential issues of reverse causality, resulting from collinearity with the wage gap, must be considered. Although I hypothesize that both female share of the workforce and state legislature impact the wage gap, rather than vice-versa, this is of course speculative, and study is limited by this possible endogeneity problem.

Limitations of the study and future implications are discussed in Chapter 5.

Chapter 5: Conclusion: Summary, Limitations, and Future Implications

The gender wage gap is entrenched in American society, and the magnitude of the gap varies across the country. To my knowledge, state gender wage gap variation has received little attention in the literature, yet there is much to gain from a deeper understanding of state wage gap variation. Throughout the thesis, I explored the question: why do different states have different gender wage gaps? With the hypothesis that a number of state level variables have potentially significant associations with state gender wage gaps, I conducted regression analysis. My findings confirm that certain variables have statistically significant relationships with state wage gaps: Hispanic population, Black population, median age, urbanization, female share of the workforce, comparable worth legislation, and women in the state legislature. I place emphasis on the variables that are significant across all models in which they are included: Hispanic population, female share of the workforce, comparable worth legislation, and women in the state legislature.

In this exploratory study, I found that the positive associations between equitable wages and urbanization, female share of workforce, and women in the state legislator all confirm my directional hypotheses (summarized in Figure 3.1) of their associations with the wage gap. I posit that the pattern in my findings of female participation, in the workforce and legislator, is important. This pattern reflects theories of the impact that greater female involvement in the workplace and policy-making has on outcomes for women. In addition to these hypothesis-confirming findings, my analysis includes findings that counter my initial hypotheses. Both the Hispanic and Black population size, and higher median ages, are positively associated with wage equality, contrary to my hypotheses, and what much of the literature implies about the national wage gap. Although I am unsure of the explanation for these results, I included the theories that

racially diverse states might have less discrimination of all types, and states with marginally higher median age demographics may have more millennials in the workforce, the age group with the lowest wage gap (Patten, 2015). Similarly, comparable worth legislation is negatively associated with wage equality, counter to my hypothesis. I suggest this is a result of the combination of liberal legislators and greater need for equal pay policy in the majority of the states. Future analysis could explore all of these findings further.

Another consideration of my findings, not discussed in my analysis, is what is not significant. No equal pay policy included in my political analysis, aside from comparable worth legislation, which is negatively related to equitable wages, is significant in the models. Given that equal pay legislation is designed to ameliorate pay discrimination, this is non-finding is troubling. I theorize that this relationship is due to the difficulty of enforcement of equal pay legislation, or relative newness of the laws, such as parental leave, which currently exists in only three states ("Paid Family Leave in Four," 2015). Regardless of why equal pay legislation is insignificant in my analysis, if only demographic factors determine the magnitude of state wage gaps, prospects for working toward equality are seemingly bleak. Perhaps future analysis can explore the effectiveness of state equal pay legislation. As public policy is the most controllable means of narrowing the wage gap, the topic of state equal pay legislation is important to study and alter in the future.

Beyond discussion of my findings, it is necessary to mention the limitations of the study, and what future research might explore. First, restrictions on the available data limited the thesis. For certain variables (urbanization, religiosity, and fertility rate) data for the year of the dependent variable, 2015, was unavailable. Although I argue that the discrepancies in the dataset caused by the year inconsistency is negligible, it is a notable limitation of my dataset. Also, the

Census Bureau reports the gender wage gap for full-time employment only. Because women disproportionately participate in the part time workforce, this potentially impacted my analysis (Patten, 2015). Another difficulty presented by data was the quantification of certain variables, particularly existing equal pay legislation. There are many types of equal pay related policies, and the content of the legislation varies. Fitting the legislation into narrowly defined dichotomous categories undoubtedly obscured some of the idiosyncrasies. Similarly, my best effort was used toward consistency in my compiled dataset, but certain challenges were presented, including: whether demographic variables should represent the entire state or strictly the working population, or if it is negligible that unionization rates are available only for the entire workforce, not restricted to full-time employees. Answers to these questions were largely subjective, and my resulting choices may have affected the findings.

The low number of observations in the study, 50 or 49 depending on the model, also limited the thesis, as a relatively low numbers of observations render few degrees of freedom, and less statistical accuracy. Unfortunately, this limitation is characteristic of studying American states, and cannot be altered for analysis of this kind. This reality forced me to limit the number of independent variables where possible. One potentially impactful decision I made to limit the number of observations was simplifying industry breakdown to only public versus private sector employment. This simplification may not have been the optimal quantification of industry, as sector differences in pay differentiation might account for a significant amount of state wage gap differentiation (Blau and Kahn, 2016). Future scholarship of state wage gap variation could benefit from more in depth analysis of state sector makeup.

Lastly, a limitation of the study is the presence of multi-collinearity in the dataset among independent variables. Although I approached the collinearity methodically, variables had to be

excluded from the models, and this may have impacted the results. Another issue of collinearity, and possibly endogeneity, was presented by the relationships of two variables to the independent variable. Although these relationships were below my threshold for collinearity, high R squared values suggest that it may have impacted some of my models. Because many of the demographic variables that are associated with wage gaps are associated with one another and the wage gap, collinearity is unfortunately characteristic of quantitative analysis of the American gender wage gap that includes demographic variables.

Although this study has limitations, many of these limitations are character of all quantitative analysis. It is my hope that despite the restrictions of quantitative analysis, the thesis has contributed to the study and discussion of the American gender wage gap. By creating a deeper understanding of the gender wage gap, society can work toward eliminating wage discrimination. Beyond pursuing a normative model of equality, closing the gender wage gap is an important and urgent goal, as women continue to play a large role in the American economy and family support. For example, "40% of families with children under 18 at home include mothers who earn the majority of the family income," and 40 percent of families are thereby hurt by gender wage discrimination ("The American Family," 2015).

Further than the gender wage gap, this study confronted the pervasiveness of gender inequality in the United States. Addressing the gender wage gap is a place to start, but parity will require more wide spanning efforts. However, progress provides hope for a more equitable future. To name a few accomplishments, the United States has never had a female president, but the 2016 election marked the first major party female nominee. Three out of the eight sitting Supreme Court justices in 2017 are women ("Supreme Court Justices," 2017). In the past century, the number of women in the House of Representatives has grown from zero to 20

percent, and women now exceed men in attainment of college degrees (“Women in the United States House,” 2017;“Population Estimates,” 2015). While a glass ceiling undoubtedly remains, as articulated by former First Lady, Secretary of State, and presidential nominee, Hillary Clinton, it has “about 18 million cracks in it, and the light is shining through like never before” (“Text of Clinton’s,” 2008).

Appendix A: Data Sources and Definitions

This is a more in depth discussion of the data and data sources used in the study.

1) **Gender wage gap:** median yearly earnings ratio for full-time, year-round workers above the age of 16, in 2015, from an American Association of University Women (AAUW) study that used American Community Survey (ACS) data ("The Gender Pay Gap by State," 2015). Although the AAUW is an organization with goal of advancement of gender equality, and is therefore bias, I hold that the data compiled by the AAUW is appropriate for the thesis because it is created with U.S. Census Bureau ACS earnings data, and given the time constraints of the project, using the AAUW calculations for state wage gaps is optimal.
2) **Industry makeup:** percentage of full time employed state population that works within the public sector (combined federal, state, and local) in 2015. Created by dividing the number of people employed full time by all government functions in each state by the full time civilian workers (16 and over) in the state, from the United States Census Bureau 2015 Annual Survey of Public Employment & Payroll, and ACS five year estimates respectively ("Population Estimates," 2015). Although this variable is created with the use of two different data points from the Census Bureau, given that it is not available to my knowledge elsewhere, using two available data points from the Census Bureau was the optimal way to include this variable in the models. This variable is inversed in the study for normal distribution.
3) **Racial makeup:** percentage in each state that identifies as white alone, non-Hispanic Black, and Hispanic alone, all in 2015 and from the United States Census Bureau ACS data ("Population Estimates," 2015). It is notable that this variable is not consistent with the use of five-year estimates where available in the thesis data set. However, these variables were not available, expressed as percentages, in the five-year estimates 2011-2015 at the time of data collection. The white alone variable is cubed in the study for normal distribution, and the Black and Hispanic population variables are expressed as the natural logarithm in the study for normal distribution.
4) **Existing equal pay legislation and parental leave policy:** Three dichotomous variables for whether or not there is existing legislation in the following categories: comparable worth, defenses and remedies, and protections. The law classifications, drawn from the AAUW, represent the laws as of 2015 ("AAUW Policy Guide," 2016). In order to lag the variable to control for the lag between legislation and impact, legislation passed in 2015 and 2016 was removed, information that was provided by the NCSL (NCSL, 2015a; NCSL, n.d.). As noted for the dependent variable, despite the bias of the AAUW, it was not feasible for this project to create classifications of all state level equal pay legislation, and it is thus optimal to utilize AAUW data. Additionally, one dichotomous variable is included for whether or not the state has paid family leave in 2015, from the NCSL (NCSL, 2016a).
5) **Age composition:** median age of the state population in 2015, from U.S. Census Bureau ACS 5 year estimates, 2011-2015 ("Population Estimates," 2015). This variable is squared in the study for normal distribution.
6) **Education levels:** the percentage of state population 25 and over with a bachelor's degree or higher, from the United States Census Bureau five year estimates, 2011-2015

("Population Estimates," 2015). This variable is expressed as the natural logarithm in the study for normal distribution.

7) **Partisan composition of governorship and state legislature:** expressed with an index of partisan majority of the state house, and partisan majority of the state senate, after the elections throughout 2004-2014, where one point is assigned for every two year period with a Republican majority in either chamber. 1 indicates republican, and 0 indicates democrat, and the scores range from 0 to 12. Compiled from *Ballotpedia*, an Encyclopedia of American politics, and NCSL data ("States-Ballotpedia," 2016; NCSL, 2014c; NCSL, 2014b). The 2004-2014 year range is used, as opposed to ending in 2015 to be consistent with the dependent variable, in order to control for the lag between the election of policy makers, legislation, and potential impact on the gender wage gap. Notably, Nebraska is missing data, as it does not have a bicameral state legislature (NCSL, 2014c). For normal distribution, this variable is expressed in the study as a dichotomous variable where a 1 indicates an above median index score and a 0 indicates a below median index score.
8) **Minimum wage level:** state mandated minimum legal dollar amount per one hour of work, from U.S. Department of Labor publication of minimum wage laws by state in 2014, lagged a year from the dependent variable, from the Labor Law Center ("Minimum Wage Rates," 2017). In states without minimum wages (Alabama, Louisiana, Mississippi, South Carolina, and Tennessee) the federal minimum wage of $7.25 is assigned (NCSL, 2016b). The Labor Law Center is a for profit organization, however to my knowledge this historical information is not readily available from a government or non-partisan non for profit organization, rather these organizations publish the current minimum wage levels. Notably, this variable is not included in the regression analysis of the study due to the inability to transform it to a normal distribution.
9) **Income:** state per capita income, or average income per person, from U.S. Census Bureau ACS 5 year estimates, 2011-2015 ("Population Estimates," 2015). This variable is expressed as the natural logarithm in the study for normal distribution.
10) **Recent vote on equal pay legislation:** whether there has been a vote on equal pay legislation in the last two complete state legislative session, 2014 and 2015, where 1 represents a vote, and 0 is a lack of a vote. Drawn from the NCSL yearly roundups of proposed equal pay related legislation (NCSL, 2015a; NCSL, 2014a).
11) **Women in state legislature:** percentage of total seats in the state legislature that were held by women, as of September 2014, from NCSL published statistics (NCSL, 2015b). This variable is lagged in order to account for the delay between election of female representatives and potential effectiveness on equal pay within the legislature.
12) **Urbanization:** the percentage of the state that lived in an urban environment ("densely developed territory, [with] residential, commercial, and other non-residential urban land uses" and "of 50,000 or more people") in 2010, from U.S. Census Bureau data ("Urban and Rural," 2010). Although this variable is from an inconsistent year with the dependent variable, it is the latest Census reporting, and is presumably representative of the 2015 values. This variable is expressed as the natural logarithm in the study for normal distribution.
13) **Unionization rate:** percentage of employed people in the state that are represented by unions, in 2015, from U.S. Bureau of Labor Statistics ("Table 5.," 2015). Notably, this data reflects all of the workforce, rather than just the full time workforce. However, the

Bureau of Labor Statistics does not publish state level data for part time unionization, and part time union representation is reportedly not a significant portion of unionization ("Union Membership," 2017).

14) **Religiosity:** overall index number comprised of religious importance, frequency of prayer, worship attendance, and belief in god, expressed as percentages that range from .77 to .33 among the states in February 2016, from Pew Research Center. Although this variable is from an inconsistent year with the dependent variable, it is the closest Pew Research Center reporting to 2015, and is presumably representative of the 2015 values.
15) **Fertility rate:** state average number of children per female lifetime as of 2014, from the Center for Disease Control (CDC) vital statistics ("List of U.S. States and Territories by Fertility," 2016). Although this variable is from an inconsistent year with the dependent variable, it is the latest CDC reporting of total fertility rate by state as of data collection, and is presumably representative of the 2015 values. This variable is inversed in the study for normal distribution.
16) **Female share of workforce:** the percentage of the full time, year-round civilian workforce aged 16 and over that consists of women, from U.S. Census Bureau ACS 5 year estimates, 2011-2015 ("Population Estimates," 2015). This variable is squared in the study for normal distribution.

Appendix B: Multi- Collinearity

Detailed correlation magnitude and direction of the eight collinear relationships, prior and post transformation where applicable, is displayed in Figure B.1. As demonstrated by Figure B.1, all relationships remain collinear after transformation. Where variables are transformed, the differences in the correlation after transformations are all below .02, with the exception of the correlation between urbanization and education, presumably because both of these variables are logged. Transformation of variables to fit normal distribution evidently has little impact on the presence of collinearity in the dataset

Figure B.1: Correlations

Collinear Pair	**Correlation**
1) Fertility rate and median age 2) Fertility rate (inversed) and median age	1) -0.7693 2) 0.7578
1) Religiosity and per capita income 2) Religiosity and income (logged)	1) -0.6468 2) -0.6675
Religiosity and women in state legislator	-0.7175
1) Education and per capita income 2) Education (logged) and per capita income (logged)	1) 0.8432 2) 0.8429
1) Women in state legislature and education 2) Women in state legislature and education (logged)	1) 0.6180 2) 0.6223
1) Religiosity and education 2) Religiosity and education (logged)	1) -0.6604 2) -0.6721
1) Urbanization and Hispanic population 2) Urbanization (logged) and Hispanic population (logged)	1) 0.6137 2) 0.8153
1) Female share of workforce and Black population 2) Female share of workforce (squared) and Black population (logged)	1) 0.5989 2) 0.6097

I will also elaborate on the information in Chapter 4 on the decisions for which variables to remove from the models. Both fertility rate and religiosity are excluded from all models because these variables were not at statistically robust as the variables that they are found to be

collinear with. Fertility rate was not collected from the same year as the dependent variable, as the most recent CDC publication of this data was in 2014, rather than 2015. Although state fertility rates are unlikely to significantly differ between 2014 and 2015, the data source technically violates the cross sectional nature of this study, and I therefore argue the lesser value of it to the study as compared to the variable it is collinear with, median age. I also argue that statistical ambiguity renders religiosity less valuable to the study than the variables it is collinear with, per capita income, women in the state legislature, and education. To create the religiosity index, the Pew Research Center uses belief and worship data. Belief and worship metrics are fundamentally less quantifiable than the rest of the independent variables included, and these metrics are also based off self-reported data, which can frequently be misrepresentative. Additionally, as religiosity has three collinear relationships in the dataset, in an effort to omit the lowest number of variables in the process of collinearity correction, it is logical to omit religiosity.

Last, I favor income over education in all models, a decision made in order to consistently include women in state legislator in all political models, which is collinear with education, and therefore could not be included in models with education.

References

AAUW Policy Guide to Equal Pay in the States. (n.d.). Retrieved from http://www.aauw.org/resource/state-equal-pay-laws/

Blau, F., & Kahn, L. (2000). Gender Differences in Pay. *NATIONAL BUREAU OF ECONOMIC RESEARCH*. Retrieved from http://www.nber.org/papers/w7732.pdf

Blau, F., & Kahn, L. (2007). The Gender Pay Gap: Have Women Gone as Far as They Can? *Academy of Management Perspectives*, *21*(1), 7–23.

Blau, Francine. (2012). The Sources of the Gender Pay Gap. *The New Gilded Age: The Critical Inequality Debates of Our Time.*, 189–208.

Blau, F., & Kahn, L. (2016). THE GENDER WAGE GAP: EXTENT, TRENDS, AND EXPLANATIONS. Retrieved from http://www.nber.org/papers/w21913.pdf

Bobbitt-Zeher, D. (2007). The Gender Income Gap and the Role of Education. *Sociology of Education*, *80*(1), 1–22. https://doi.org/10.1177/003804070708000101

Browne, J. (2014). *Corporate Boards, Quotas for Women, and Political Theory*. Wolfson College, University of Oxford. Retrieved from http://www.fljs.org/sites/www.fljs.org/files/publications/Browne.pdf

Cavalcanti, T., & Tavares, J. (2016). The Output Cost of Gender Discrimination: A Model-Based Macroeconomics Estimate. *Economic Journal*, *126*(590), 109–134.

Childs, S., & Krook, M. L. (2008). Critical Mass Theory and Women's Political Representation. *Political Studies*, *56*. https://doi.org/10.1111/j.1467-9248.2007.00712.x

Employed persons by sex, occupation, class of worker, full- or part-time status, and race. (2017). Retrieved February 22, 2017, from https://www.bls.gov/cps/cpsaat12.htm

Erosa, A., Fuster, L., & Restuccia, D. (2016). A quantitative theory of the gender gap in wages.

European Economic Review, *85*, 165–187. https://doi.org/10.1016/j.euroecorev.2015.12.014

Hoffman, S. D. (2013). Geography And Gender: Why Does the Gender Earnings Ratio Vary Across U.S. States? Retrieved October 22, 2016, from http://www.econbiz.de/Record/geography-and-gender-why-does-the-gender-earnings-ratio-vary-across-u-s-states-hoffman-saul/10010815435/Description

Jarrell, Stephen, S., T. D. (2004). Declining Bias and Gender Wage Discrimination? A Meta-Regression Analysis. *The Journal of Human Resources (JHR)*, *XXXIX*(3), 828–838. https://doi.org/10.3368/jhr.XXXIX.3.828

Johnson, G., & Solon, G. (1986). Estimates of the Direct Effects of Comparable Worth Policy. *The American Economic Review*, *76*(5), 1117–1125.

Johnson, J., Reynolds, H. T., & Mycoff, J. (2015). *Political Science Research Methods*. Sage Publications.

Jung, H., & Cho, J. (2014). Quality of Jobs for Female Workers: A Comparative Study of South Korea and Australia. *Applied Research in Quality of Life*, *11*(1), 1–22. https://doi.org/10.1007/s11482-014-9351-1

Kellstedt, P., & Whitten, G. (2009). *The Fundamentals of Political Science Research* (1st ed.). Cambridge University Press;

Kerr, S. P. (2016). Parental Leave Legislation and Women's Work: A Story of Unequal Opportunities. *Journal of Policy Analysis and Management*, *35*(1), 117–144. https://doi.org/10.1002/pam.21875

Kim, Marlene. (2013). Policies to End the Gender Wage Gap in the United States. *Review of Radical Political Economics*, *45*, 278–283. https://doi.org/10.1177/0486613413487159

Lipka, M., & comments, B. W. (2016, February 29). How religious is your state? Retrieved from http://www.pewresearch.org/fact-tank/2016/02/29/how-religious-is-your-state/

List of U.S. states and territories by fertility rate. (2016, August 27). In *Wikipedia*. Retrieved from https://en.wikipedia.org/w/index.php?title=List_of_U.S._states_and_territories_by_fertility_rate&oldid=736421189

List of women CEOs of Fortune 500 companies. (2016, September 1). In *Wikipedia, the free encyclopedia*. Retrieved from https://en.wikipedia.org/w/index.php?title=List_of_women_CEOs_of_Fortune_500_companies&oldid=737285243

Livingston, G. (2016, September 26). Among 41 nations, U.S. is the outlier when it comes to paid parental leave. Retrieved from http://www.pewresearch.org/fact-tank/2016/09/26/u-s-lacks-mandated-paid-parental-leave/

Mason, P. L. (2010). Moments of Disparate Peaks: Race-Gender Wage Gaps Among Mature Persons, 1965–2007. *The Review of Black Political Economy*, *38*(1), 1–25. https://doi.org/10.1007/s12114-010-9059-x

Measuring the Wage Gap Archives. (n.d.). Retrieved from https://nwlc.org/issue/measuring-the-wage-gap/

Median weekly earnings by educational attainment in 2014 : The Economics Daily: U.S. Bureau of Labor Statistics. (2014). Retrieved January 24, 2017, from https://www.bls.gov/opub/ted/2015/median-weekly-earnings-by-education-gender-race-and-ethnicity-in-2014.htm

Minimum Wage Rates by State. (2017). Retrieved January 27, 2017, from

https://www.laborlawcenter.com/state-minimum-wage-rates/

National Conference of State Legislatures. (2014a). *2014 Equal Pay Legislation*. Retrieved from http://www.ncsl.org/research/labor-and-employment/2014-equal-pay-legislation.aspx

National Conference of State Legislatures. (2014b). *2014 State and Legislative Partisan Composition*. Retrieved from http://www.ncsl.org/Portals/1/Documents/Elections/Legis_Control_2014_Dec2_11am.pdf

National Conference of State Legislatures. (2014c). *Partisan composition of State Legislatures 2002-2014*. Retrieved from http://www.ncsl.org/documents/statevote/legiscontrol_2002_2014.pdf

National Conference of State Legislatures. (2015a). *2015 Equal Pay Legislation*. Retrieved from http://www.ncsl.org/research/labor-and-employment/equal-pay-laws.aspx

National Conference of State Legislatures. (2015b). *Women in State Legislatures for 2014*. Retrieved from http://www.ncsl.org/legislators-staff/legislators/womens-legislative-network/women-in-state-legislatures-for-2015.aspx

National Conference of State Legislatures. (2016a). *State Family and Medical Leave Laws*. Retrieved from http://www.ncsl.org/research/labor-and-employment/state-family-and-medical-leave-laws.aspx

National Conference of State Legislatures. (2016b). *State Minimum Wages| 2016 Minimum Wage by State*. Retrieved from http://www.ncsl.org/research/labor-and-employment/state-minimum-wage-chart.aspx

National Conference of State Legislatures. (n.d.). *State Equal Pay Laws – August 2016*. Retrieved from http://www.ncsl.org/research/labor-and-employment/equal-pay-laws.aspx

Patten, E. (2015, April 14). On Equal Pay Day, key facts about the gender pay gap. Retrieved from http://www.pewresearch.org/fact-tank/2015/04/14/on-equal-pay-day-everything-you-need-to-know-about-the-gender-pay-gap/

Pew Research Center. (2015). *The American Family Today*. Retrieved from http://www.pewsocialtrends.org/2015/12/17/1-the-american-family-today/

Polachek, S. W. (2014). Equal pay legislation and the gender wage gap. *IZA World of Labor*. https://doi.org/10.15185/izawol.16

Population estimates, July 1, 2015, (V2015). (n.d.). Retrieved November 22, 2016, from //www.census.gov/quickfacts/

Robust Regression | Stata Data Analysis Examples. (n.d.). Retrieved from http://stats.idre.ucla.edu/stata/dae/robust-regression/

Rutgers University Eagleton Institute of Politics. (2015). *Women in U.S. Congress 2015*. Retrieved from http://www.cawp.rutgers.edu/women-us-congress-2015

Ryu, K. (2010). STATE POLICIES AND GENDER EARNINGS INEQUALITY: A Multilevel Analysis of 50 U.S. States Based on U.S. Census 2000 Data. *Sociological Quarterly*, *51*(2), 226–254. https://doi.org/10.1111/j.1533-8525.2010.01170.x

Sebawit, B. G., & Alkadry, M. (2016). A Systematic Review of the Gender Pay Gap and Factors That Predict It. *Administration & Society*. https://doi.org/10.1177/0095399716636928

Seguino, S. (2011). Help or Hindrance? Religion's Impact on Gender Inequality in Attitudes and Outcomes. *World Development*, *39*(8), 1308–1321. https://doi.org/10.1016/j.worlddev.2010.12.004

States - Ballotpedia. (n.d.). Retrieved December 14, 2016, from https://ballotpedia.org/States

Street, 1615 L., NW, Washington, S. 800, & Inquiries, D. 20036 202 419 4300 | M. 202

419 4349 | F. 202 419 4372 | M. (2013, December 11). On Pay Gap, Millennial Women Near Parity – For Now. Retrieved from http://www.pewsocialtrends.org/2013/12/11/on-pay-gap-millennial-women-near-parity-for-now/

Study Finds the Gender Wage Gap Is Larger for Black Women Than It Is for White Women. (2013). *Journal of Blacks in Higher Education (Online)*. Retrieved from http://search.proquest.com/docview/1268836052/abstract/462F5A8F43D44AA3PQ/1

Supreme Court of the United States. (2017). *Biographies of Current Justices of the Supreme Court*. Retrieved from https://www.supremecourt.gov/about/biographies.aspx

Table 5. Union affiliation of employed wage and salary workers by state. (2015). Retrieved November 1, 2016, from http://www.bls.gov/news.release/union2.t05.htm

Text of Clinton's 2008 concession speech. (2008, June 7). *The Guardian*. Retrieved from https://www.theguardian.com/commentisfree/2008/jun/07/hillaryclinton.uselections2008 1

Tharenou, P. (2012). The Work of Feminists is Not Yet Done: The Gender Pay Gap—a Stubborn Anachronism. *Sex Roles*, *68*(3–4), 198–206. https://doi.org/10.1007/s11199-012-0221-8

The Gender Pay Gap by State and Congressional District. (2015). Retrieved from http://www.aauw.org/resource/gender-pay-gap-by-state-and-congressional-district/

United States Government Accountability Office. (2011). *Gender Pay Differences: Progress MAde, but Women Remain Overrepresented among Low-Wage Workers* (Report to Congressional Requesters).

Union Membership (Annual) News Release. (2017). Retrieved April 4, 2017, from https://www.bls.gov/news.release/union2.htm

Urban and Rural Population by State, 2010. (2010). Retrieved from

http://webcache.googleusercontent.com/search?q=cache:6Bx23AkSii4J:cber.cba.ua.edu/edata/census2010/Urban%2520Rural%2520by%2520State%25202010%2520short%2520ver.xls+&cd=4&hl=en&ct=clnk&gl=us&client=safari

U.S. Census Bureau. (n.d.). FAQs. Retrieved October 18, 2016, from https://www.census.gov/topics/income-poverty/income/about/faqs.html

U.S. Department of Labor. (1963). *American Women: Report of the President's Commission on the Status of Women*. Retrieved from https://www.dol.gov/wb/American%20Women%20Report.pdf

Weichselbaumer, Doris, W.-E., Rudolf. (2005). A Meta-Analysis of the International Gender Wage Gap. *Journal of Economics Surveys*, *19*(3), 479–511.

Wolbrecht, C. (2000). *The Politics of Women's Rights: Parties, Positions, and Change*. Princeton University Press. Retrieved from http://www.jstor.org/stable/j.ctt7t299

Wolfe, J., & Fields, E. (1995). Interindurstry Wage Differentials and the Gender Wage Gap. *SAGE Journals*. https://doi.org/doi: 10.1177/001979399504900107

Women in the United States House of Representatives. (2017, March 23). In *Wikipedia*. Retrieved from https://en.wikipedia.org/w/index.php?title=Women_in_the_United_States_House_of_Representatives&oldid=771760823