A Study Analyzing the Impact of Unionization on State Wages

Mckenzie Rhone, Sarah Tinsley, Ellie Tsuchiya

ECON 3161: Econometric Analysis

Dr. Shatakshee Dhongde

Fall 2018

Abstract:

This study was motivated by existing literature on unions which attempt to understand and measure the impact of unionization on wages and the welfare of workers. Our analysis tries to further this research by exploring the effects unionization has on wages at the state level. Unionization, our main explanatory variable, is identified as the total union coverage by state in both the private and public sectors in all of our models. The explained variable in all of our models is mean hourly wages. Our results suggest that union coverage positively impacts wages. Furthermore, the size of the IT professional sector within a given state's economy has the largest impact on hourly wages followed by union coverage and GDP.

I. Introduction

The effect of unionization on the labor market and the welfare of workers has been a longstanding question in social science research. Much attention has been dedicated to understanding the effects of unionization on wages, specifically wage differentials between covered workers and uncovered workers. Unions act as a collective bargaining method for workers and have both monopoly and voice effects (Freeman & Medoff, 1985). Workers ban together as a "single seller" of labor and also push for benefits beyond monetary compensation. Unions help ensure that some of the economic rents of the company are reallocated to the workers. It is important to understand the effects of unionization on wages in order to measure the effectiveness of collective bargaining in the workplace and to determine whether this type of labor structure is truly helpful.

Despite a rich literature on unionization, it is inherently difficult to isolate a causal relationship between unions and wages. This is partly due to the difficulty of observing a counterfactual, the fact that unionization is not randomly assigned, and the high likelihood of biased estimators. Factors which affect wages also likely affect worker and employer selection into unions. Unions will be more appealing to workers who face a higher probability of lower earnings, which raises the demand for union jobs and allows employers to be more selective on awarding union positions. Consequently, worker selection implies a negative bias and employer selection implies a positive bias on the estimated wage gap (Farber, 2001).

While we do not have access to the ideal dataset, we are seeking to identify a relationship between unionization and wages at the state level in the U.S. We hypothesize that, ceteris paribus, states with a higher percentage of unionization will, on average, have higher mean wages. We predict this relationship holds because of the monopoly power unions hold and the policy protections (e.g., the National Labor Relations Act) that prevent employers from prohibiting unionization. Unions are a mechanism for collective action, norm creation in the workplace, and can stimulate higher wages on average in the nonunion sector through threat effects.

II. Literature Review

Although the first union in the U.S. materialized in 1881, unionization really took off post-depression and grew in strength and numbers during the 40s and 50s. Since then, however, union membership has declined significantly. According to the 2017 Bureau of Labor Statistics, in 1983, the union membership rate in the U.S. was 20.1 percent. In 2017, it was 10.7 percent. From 1973 to 2007, union membership in the private sector declined from 34 to 8 percent for men and from 16 to 6 percent for women (Western & Rosenfeld, 2011). Yet, despite the declining prevalence of unions, they still prove to be influential in the U.S. labor market and in politics.

A common approach to studying the effects of unionization on wages is to observe the wage differential between unionized and non-unionized workers by industry or to use longitudinal data to measure the wage gains and losses to workers who change union status.

The work of Gregg Lewis, perhaps one of the most influential economists in labor economics, was pioneering in regards to studying the wage effects of unions (Sherwin, 1994). He published two books on the topic—one in 1963 and another in 1986. In his works, Lewis studied the impact of unionization on relative wages in the U.S. economy spanning the years of 1920-1970 using aggregated data at the industry level (Sherwin, 1994). He divided the economy into two sectors and, after controlling for a variety of factors, compared the movement in the log of relative wages with the difference in the fraction of workers unionized. Importantly, he distinguished between "wage gaps" and "wage gains," with the former being the difference in wages between union members and nonmembers and the latter representing the wage effects of unionism on union wages in comparison to what wages would have been without unionization. As aforementioned in our introduction, empirically studying the latter is difficult because we cannot observe what wages would have been in the same situation without unionization. Thus, research in this area has focused on estimating wage gaps/wage differentials between workers covered by unions and workers who are not. Lewis concluded that from 1967-1979 the U.S. mean wage gap was approximately 15 percent (Blanchflower, David and Bryson, 2004) and that this wage gap varied depending on sector and worker characteristics.

Other research has corroborated the positive effect of unionization on wage gaps for union members. Pencavel and Hartsog's work in 1984, "A Reconsideration of the Effects of Unionism on Relative Wages and Employment in the US 1920-1980," found strong evidence of a positive impact of unionization on the average wage of union workers relative to nonunion workers. Improvements in social welfare can be made if equity gains, in the form of distribution of income, outweigh efficiency losses due to unemployment (Oswald, 1982). Additionally, he found the wage gap was largest for workers with lower levels of observed skills.

Very little recent literature, though, explores wage effects of unionization in this new era of declining union popularity in the United States. Bruce Western and Jake Rosenfeld sought to address this gap and determine if rising wage inequality in the U.S. is related to the decline in union membership. Their analysis discovered that a decline in unionization has led to an increase in wage inequality explaining a fifth to a third of the growth in inequality. They argued that unions helped institutionalize norms of equity which reduced the variance of nonunion wages in highly unionized industries and regions (Western and Rosenfeld, 2011). Since unionization is declining, these positive externalities of unions are no longer materializing.

The lack of updated research on union wage effects represents a gap in the literature and an avenue for further academic exploration. Our paper helps add to this gap by exploring union wage effects using recent data. Additionally, most studies use micro data at the firm level to determine the impact of unionization on wage gaps. We are opting to explore the effects of unionization at the state level. Instead of comparing wages between union and nonunion members, we will be comparing wages at the state level between states with varying unionization strength and presence. If unions have a strong impact on wages then wage differences across states could be explained by the varying levels of unionization at the state level.

III. Data

In order to analyze the ceteris paribus effects of unionization on wage, we have selected a single dependent variable and relevant independent variables for our various regression models. These variables were chosen based on conclusions drawn from our research which indicated that they could show both historically and statistically significant relationships between wage and unionization. Due to limitations in data, for each variable, we have compiled the most current data available across a variety of trusted sources. We use state-level data excluding Washington D.C. which proved to be an outlier with the highest mean hourly wage in the United States.

Dependent Variable

Mean Hourly Wage*, coded as meanhourwage*

Unions play a large role in raising wages for workers through collective bargaining practices. Numerous economic studies have shown that unionized workers earnings exceed those of nonunion workers, creating a "union wage premium". Historically, unions have lessened wage inequality by giving larger premiums to low-skilled workers than for high-skilled workers. Ideally, we would have used mean hourly wages by state, by industry to control for wage differences within industries. However, due to limitations in data, we use mean hourly wage data by state across all occupations in 2017 as reported by the Bureau of Labor Statistics.

Independent Variables

Union Coverage*, coded as percenttotalunioncov*

Unions must equally represent all covered employees regardless of their membership status. A covered non-member does not pay membership fees and therefore does not have decision making rights in the union. However, they reap the same benefits as a member from being represented by a union. Total union coverage across both public and private sectors is used rather than union membership to avoid a biased estimate of the union/nonunion differential for select groups of workers. This variable is measured by the Current Population Survey of 2017, which calculates the number of employees who report being covered by a union divided by the number of employed people surveyed, expressed as a percentage. The values for this variable are coded as percentages (between 0 and 100), as collected from the corresponding data source.

Real GDP per capita, *coded as log_realgdppercapita*

Gross domestic product (GDP) per capita is calculated by dividing the total economic output by the population with adjustments for inflation. When GDP is high, production is increasing, and consumers have more disposable income and more incentive to spend. Firms are willing to hire more workers and pay higher wages in periods of economic prosperity. Therefore, we expect to see a positive relationship between GDP and mean hourly wages. We take the logarithm of this variable in order to normalize the values and results since there is such a large range of values for the GDP per capita variable. The data used in our model is from the Bureau of Economic Analysis (2018).

Unemployment rate*, coded as unemploymentrate*

The 2018 unemployment rate by state from the Bureau of Labor Statistics is included as a variable to account for the macroeconomic conditions of a given state. In a competitive labor market, wages are determined by supply and demand. In times of high unemployment, the ample supply of labor disincentivizes companies to offer competitive wages which could be reflected in the mean hourly wage. On the other hand, in times of low unemployment, the scarce supply of labor incentivizes companies to set higher wages to attract the best talent which could be reflected in the mean hourly wage.The values for this variable are coded as percentages (between 0 and 100), as collected from the corresponding data source.

Public high school graduation rate*, coded as gradratehs*

The Human Capital Theory states that the accumulation of human capital (knowledge, habits, social and personality characteristics) increases worker productivity. We use the public high school graduation rate by state to determine the effects completing high school may have on wages. This data is from the National Center for Education Statistics for 2015-2016. The values for this variable are coded as percentages (between 0 and 100), as collected from the corresponding data source.

Percent manufacturing*, coded as percentmanufacturing*

As previously stated, we were not able to find mean hourly wage data by state, by industry to account for the wage differentials within industries. We attempt to control for these wage differentials by introducing a variable that captures the percentage of employees in the manufacturing sector by state. We predict that a greater percentage of employees in the manufacturing sector will reflect lower wages due to the minimal skill level required for this job. The values for this variable are coded as decimals (between 0 and 1), as calculated from the corresponding data source.

Percent IT, *coded as percentitprofessionalservices*

As previously stated, we were not able to find mean hourly wage data by state, by industry to account for the wage differentials within industries. We attempt to control for these wage differentials by introducing a variable that captures the percentage of employees in the IT sector by state. We predict that a greater percentage of employees in the IT sector will reflect higher hourly wages due to the higher level of education required and high demand for these jobs. The values for this variable are coded as decimals (between 0 and 1), as calculated from the corresponding data source.

Percent graduate degree, *coded as percentgraddegree*

In line with the Human Theory Capital, we predict that higher education in the form of a graduate or professional degree will have an impact on mean hourly wages. Therefore, we include the percent of adults 25-64 with a graduate or professional degree by state to our regression analysis. This data is from the 2015 American Community Survey conducted by the National Center for Higher Education Management Information Center. The values for this variable are coded as percentages (between 0 and 100), as collected from the corresponding data source.

Table 1. Variable Summary.

Variable	Description	Source	Year
meanhourwage	real average hourly wage by state	Bureau of Labor Statistics	2017
percenttotalunioncov	percentage of employed workers who are covered by a collective bargaining agreement; (Covered / Employment) * 100	Current Population Survey	2017
log_realgdppercapita	log of real GDP by state; real personal income divided by midyear population	Bureau of Economic Analysis	2016
unemploymentrate	percentage of unemployed by state	Bureau of Labor Statistics	2018
gradratehs	public high school graduation rate by state	National Center for Education Statistics	2015
percentmanufacturing	percentage of employees in manufacturing sector by state	U.S. Census Bureau	2015
percentitprofessionalservices	percentage of employees in IT sector by state	U.S. Census Bureau	2015
percentgraddegree	percentage of working adults 25-64 with a graduate or professional degree by state	National Center for Higher Education Management Information Center	2015

Summary Statistics

Below is a table of the descriptive statistics of our dependent and independent variables.

Table 2. Summary Statistics.

Descriptive Statistics					
Variable	**Observation**	**Mean**	**St.dev**	**Min**	**Max**
meanhourwage	50	23.26	2.80	18.71 (Mississippi)	29.86 (Massachusetts)
percenttotalunioncov	50	11.19	5.17	3.9 (South Carolina)	25.3 (New York)
log_realgdppercapita	50	47401.06	.09	10.59 (Mississippi)	10.97 (Massachusetts)
unemploymentrate	50	3.74	0.81	2.2 (Hawaii)	5.5 (Louisiana)
gradratehs	50	84.14	4.58	71 (New Mexico)	91 (Iowa)
percentmanufacturing	50	0.09	0.04	0.02 (Hawaii)	0.17 (Indiana)
percentitprofessionalservices	50	0.06	0.02	0.00 (West Virginia)	0.12 (Virginia)
percentgraddegree	50	11.03	2.77	7.52 (Nebraska)	18.91 (Maryland)

Gauss Markov Assumptions

1. Linear in Parameters

All of the parameters in all of our models are linear, as demonstrated below.

Simple regression model:

meanhourwage = $\beta_0 + \beta_1$percenttotalunioncov + u

Multiple regression models:

Full MLR model:

meanhourwage = $\beta_0 + \beta_1$percenttotalunioncov + $\beta_2$log_realgdppercapita + $\beta_3$unemploymentrate + $\beta_4$gradratehs + $\beta_5$percentmanufacturing + $\beta_6$percentitprofessionalservices + $\beta_7$percentgraddegree + u

Modified MLR model:

meanhourwage = $\beta_0 + \beta_1$percenttotalunioncov + $\beta_2$log_realgdppercapita + $\beta_3$percentitprofessionalservices + u

2. Random Sampling

We conduct our analysis under the assumption that all of our sources have collected their data using random sampling techniques.

3. SLR: sample variation in the explanatory variable

The sample outcomes on x, namely {xi : i = 1, 2, ..., n}, are not all the same value.

MLR: no perfect collinearity

Table 3. Correlation Between Independent Variables.

Correlation Coefficients							
Variable	**percenttotalunioncov**	**log_realgdppercapita**	**unemploymentrate**	**gradratehs**	**percentmanufacturing**	**percentitprofessionalservices**	**percentgraddegree**
percenttotalunioncov	1.0000						
log_realgdppercapita	0.2583	1.0000					
unemploymentrate	0.0736	–0.2456	1.0000				
gradratehs	–0.1288	0.2756	–0.3569	1.0000			
percentmanufacturing	–0.2264	–0.0370	–0.2101	0.3852	1.0000		
percentitprofessionalservices	0.1288	0.3409	0.0588	–0.0680	–0.3263	1.0000	
percentgraddegree	0.0659	–0.0161	–0.0235	–0.0905	–0.3323	0.3123	1.0000

4. Zero conditional mean

The expected error value to pass this assumption is 0. By including multiple independent variables, we try to make our expected error as close to 0 as possible. In order to adjust for state level socio-economic conditions, the current multiple regression models include variables such as unemployment rate, GDP, and minimum wage. We take into account the size of the workforce, as that may have an indirect effect on employee pension contributions.

However, one problem we encounter in all models is the wage differentials among industries. Although we include two variables to control for the percentage of high paying sectors in each state (IT sector) and low paying sectors in each state (manufacturing sector), this is an oversimplified solution to the problem. We could not find sector/occupational union and wage data by state to try to avoid this bias.

5. Homoscedasticity (Refer to **Figures 3 & 4** in the **Appendix** for the residual plots)

The residual plots show no significant pattern of the residuals plotted against the fitted values in the SLR Model, Full MLR Model, and Modified MLR Model, which suggests homoscedasticity for each of our models.

Satisfaction of assumptions 1 through 4 implies that our beta coefficient estimates are unbiased. Satisfaction of the 5th assumptions allows us utilize the OLS variance formulas for analysis.

IV. Results

Table 4. Model Results.

OLS Coefficient Results — Dependent Variable: meanhourwage			
Independent Variables	**SLR Model** df = 49 (c.v. 10% = 1.68) (c.v. 5% = 2.01) (c.v. 1% = 2.69)	**Full MLR Model** df = 42 (c.v. 10% = 1.68) (c.v. 5% = 2.01) (c.v. 1% = 2.69)	**Modified MLR Model** df = 46 (c.v. 10% = 1.68) (c.v. @ 5% 2.01) (c.v. @ 1% 2.69)
percenttotalunioncov	0.3547*** (0.0592)	0.2653 *** (6.48)	0.2852 *** (7.13)
log_realgdppercapita	—	9.6839 *** (3.75)	8.3838*** (3.55)
unemploymentrate	—	0.2978 (1.11)	—
gradratehs	—	0.0216 (0.43)	—
percentmanufacturing	—	–10.2114 (–1.53)	—
percentitprofessionalservices	—	48.6226 *** (4.43)	58.6364 *** (5.82)
percentgraddegree	—	0.0669 (0.85)	—
intercept	19.2919 (0.7281)	–89.4499 (–3.31)	–73.4520 (–2.92)

*** significant at 1%, ** significant at 5%, * significant at 10%
NOTE: values in the parenthesis are t-statistics

Simple Linear Regression Model (Refer to **Figure 5** in the **Appendix** for the STATA output)

population: meanhourwage = $\beta_0 + \beta_1$percenttotalunioncov + u

sample: meanhourwage = 19.2919 + 0.3547*percenttotalunioncov* + u

In our simple linear regression, we looked at how the percent of union coverage by state has an impact on the mean hourly wage. We found union coverage to have a positive and statistically significant impact on mean hourly wage. As the percent of union coverage increased by 1 percentage point, mean hourly wage increased by approximately 35 cents. It is important to note that we did not take the log of wage because the variation in mean hourly wage was small.

Multiple Linear Regression Model

Full MLR Model: (Refer to **Figure 6** in the **Appendix** for the STATA output)

population: meanhourwage = $\beta_0 + \beta_1$percenttotalunioncov + $\beta_2$log_realgdppercapita + $\beta_3$unemploymentrate + $\beta_4$gradratehs + $\beta_5$percentmanufacturing + $\beta_6$percentitprofessionalservices + $\beta_7$percentgraddegree + u

sample: meanhourwage = –89.4499 + 0.2653*percenttotalunioncov* + 9.6839*log_realgdppercapita* + 0.2978*unemploymentrate* + 0.0216*gradratehs* – 10.2114*percentmanufacturing* + 48.6226*percentitprofessionalservices* + 0.0669*percentgraddegree*

In order to conduct a more comprehensive analysis and to decrease the probability of omitted variable bias, we incorporated more variables into our regression. We included GDP and unemployment to act as indicators of the macroeconomic environment at the state level. We also included high school graduation rate and the percent of adults with a graduate degree or higher to capture the education and skill level of the state's workforce. Lastly, we added two variables, percentage of the sector in manufacturing and percentage in IT professional services, to help adjust for the variation in sector composition by state.

According to our model, only percent union coverage, GDP, and percent of IT professional sector had an impact on mean hourly wages. As union coverage increased by 1 percentage point, mean hourly wage increased by 0.2653 dollars. As GDP increased by 1 percent, mean hourly wage increased by 0.0968 dollars. As the percent of employees in the IT professional sector in the economy increased by 1 percentage point, mean hourly wage increased by 0.4862 dollars (48.626/100). The other variables were not significant at the 10 percent level. The effect of union coverage on wage decreased from our simple regression model to our multiple regression model as a result of the strength of the coefficients in our added variables. The direction of the effects of GDP, union coverage, and IT sector are in line with our predictions. The relatively large coefficient of the IT sector on wages indicates a strong need to improve our regression by creating an industry index to account for the percentage of the economy that is controlled for by above average/high paying industries. Because we could not find this data, we tried to account for the effect of high paying versus low paying jobs by using manufacturing and IT as proxies; however, this is is an incomplete method. Manufacturing showed no significant effect on wages despite our prediction that it would because it is a low-paying sector. However, manufacturing might not have been the best sector to choose as a proxy for the composition of low-paying jobs in the state.

Surprisingly, our education variables had no significant effect on wages. A more educated workforce would, theoretically, lead to a more productive workforce. Productivity is a component of economic growth which, as we have seen, impacts wages. Additionally, it is well documented that workers with more education have higher-paying jobs, on average. Thus, we expected education to have a positive and significant effect. To explore this further, we ran joint significant tests on education in our **Extensions** section.

Unemployment also had no significant effect on wages. As wages are sticky in the short run and our analysis uses cross-sectional data, we did not have strong expectations on how unemployment would impact wages across states. From the Phillips Curve, we know that the lower the unemployment rate, the tighter the labor market and the faster firms must raise wages to attract workers. On the opposite end, the higher the unemployment rate, the slower wages will grow. However, this analysis necessitates the use of time series data. To explore unemployment's effect on wages further, we ran joint significance tests on unemployment in our **Extensions** section.

One should note that for our MLR models, the interpretation of the intercept does not have much meaning, as it implies that every other variable in the model is 0, which is unrealistic.

Table 5. Classical T-Test and Other Significance Tests.

OLS Coefficient Results — Dependent Variable: meanhourwage **NOTE: these values are for the 5% significance level			
Independent Variables	**P > | t |**	**T-Stat**	**Confidence Interval**
percenttotalunioncov ***	0.000	6.48	[0.1827, 0.3479]
log_realgdppercapita ***	0.001	3.75	[4.4765, 14.8914]
unemploymentrate	0.275	1.11	[–0.2455, 0.8411]
gradratehs	0.673	0.43	[–0.0811, 0.1243]
percentmanufacturing	0.133	–1.53	[–23.6700, 3.2473]
percentitprofessionalservices ***	0.000	4.43	[26.4950, 70.7502]
percentgraddegree	0.400	0.85	[–0.0918, 0.2256]

For each explanatory variable, the tested null hypothesis was H_0: $\beta = 0$. Our degrees of freedom ranged from 42 to 49—because this is such a small range, the critical values were the same for each model—and the critical values for our t-tests were: 1.68 at the 10% level, 2.01 at the 5% level, and 2.69 at the 1% level. Of all the values of our explanatory variables, only three—percenttotalunioncov, log_realgdppercapita, and percentitprofessionalservices—yielded values greater than the critical values, allowing us to reject the null hypothesis. When looking at the p-values and confidence intervals, those tests yielded the same results as the classical t-test, with only those same three variables having values satisfying the 1%, 5%, and 10% levels of significance. Thus, out of our seven chosen explanatory variables, only the percenttotalunioncov, log_realgdppercapita, and percentitprofessionalservices proved to be significant, all three of which were at significant at the 1% level.

Modified MLR Model: (Refer to **Figure 7** in the **Appendix** for the STATA output)

population: meanhourwage = $\beta_0 + \beta_1$percenttotalunioncov + $\beta_2$log_realgdppercapita + $\beta_3$percentitprofessionalservices + u

sample: meanhourwage = –73.4520 + 0.2852*percenttotalunioncov* + 8.3838*log_realgdppercapita* + 58.6364*percentitprofessionalservices*

In our modified model, we dropped the variables from our Full MLR Model that were insignificant. We found the coefficients of our significant variables to be: 58.6364*** for IT sector, 8.3838*** for logGDP, and 0.2852*** for union coverage.

V. Extensions/Robustness Checks

We chose to run the following robustness checks in order to investigate whether certain variables that were independently insignificant might still have an effect on wages when combined with other variables. Specifically, we wanted to further investigate the role of education and unemployment on wages. Our tests for joint significance showed us that the percent of the state's population with a graduate degree or higher and the percentage of the state's economy captured by the IT professional services sector are jointly significant. Additionally, high school graduation rate and GDP were jointly significant. Unemployment and union coverage were jointly significant, and unemployment and GDP were also jointly significant.

Robustness Checks: F-Tests

Unrestricted Model: (Refer to **Figure 6** in the **Appendix** for the STATA output)

meanhourwage = $\beta_0 + \beta_1$percenttotalunioncov + $\beta_2$log_realgdppercapita + $\beta_3$unemploymentrate + $\beta_4$gradratehs + $\beta_5$percentmanufacturing + $\beta_6$percentitprofessionalservices + $\beta_7$percentgraddegree + u

****NOTE:** For the restricted models, for the purpose of this paper, the beta coefficient subscripts are left as the same numbers for easier identification of which variables were dropped.

Testing Impact of Education via Joint Significance:

1) **F-test: percentitprofessionalservices and percentgraddegree**

Restricted Model: (Refer to **Figure 8** in the **Appendix** for the STATA output)

population: meanhourwage = $\beta_0 + \beta_1$percenttotalunioncov + $\beta_2$log_realgdppercapita + $\beta_3$unemploymentrate + $\beta_4$gradratehs + $\beta_5$percentmanufacturing + u

sample: meanhourwage = –127.5642 + 0.2571*percenttotalunioncov* + 2.9255*log_realgdppercapita* + 0.3737*unemploymentrate* + 0.0151*gradratehs* – 20.5692*percentmanufacturing*

H_0: $\beta_6 = \beta_7 = 0$	**H_a = H_0 not true**
Unrestricted SSR = 79.5085477	Restricted SSR = 125.287748
df = 44	q = 2

5% critical value = 3.21 1% critical value = 5.12

F-stat = [(125.287748 – 79.5085477) / 79.5085477] * (44 / 2) = **12.6668**

We reject the null hypothesis at both the 5% and 1% significant levels, which indicates that the variables for the percentage of IT professional services sector and the percentage of graduate degrees are jointly significant.

2) F-test: gradratehs and percentgraddegree and log_realgdppercapita

Restricted Model: (Refer to **Figure 9** in the **Appendix** for the STATA output)

population: meanhourwage = $\beta_0 + \beta_1$percenttotalunioncov + $\beta_3$unemploymentrate + $\beta_5$percentmanufacturing + $\beta_6$percentitprofessionalservices + u

sample: meanhourwage = 17.0985 + 0.3071*percenttotalunioncov* – 0.0558*unemploymentrate* + –8.7978*percentmanufacturing* + 65.7374*percentitprofessionalservices*

H_0: $\beta_2 = \beta_4 = \beta_7 = 0$ **H_a = H_0 not true**

Unrestricted SSR = 79.5085477 Restricted SSR = 110.609198

df = 45 q = 3

5% critical value = 3.20 1% critical value = 5.11

F-stat = [(110.609198 – 79.5085477) / 79.5085477] * (45 / 3) = **5.8674**

We reject the null hypothesis at both the 5% and 1% significant levels, which indicates that the variables for the percentage of high school graduates, the percentage of graduate degrees, and the log of the real GDP per capita are jointly significant.

3) F-test: gradratehs and log_realgdppercapita

Restricted Model: (Refer to **Figure 10** in the **Appendix** for the STATA output)

population: meanhourwage = $\beta_0 + \beta_1$percenttotalunioncov + $\beta_3$unemploymentrate + $\beta_5$percentmanufacturing + $\beta_6$percentitprofessionalservices + $\beta_7$percentgraddegree + u

sample: meanhourwage = 16.8539 + 0.3073*percenttotalunioncov* – 0.0495*unemploymentrate* – 8.3699*percentmanufacturing* + 65.1547*percentitprofessionalservices* + 0.0194*percentgraddegree*

H_0: $\beta_2 = \beta_4 = 0$ **H_a = H_0 not true**

Unrestricted SSR = 79.5085477 Restricted SSR = 110.490475

df = 44 q = 2

5% critical value = 3.21 1% critical value = 5.12

F-stat = [(110.490475 – 79.5085477) / 79.5085477] * (44 / 2) = **8.5727**

We reject the null hypothesis at both the 5% and 1% significant levels, which indicates that the variables for the percentage of high school graduates and the log of the real GDP per capita are jointly significant.

Testing Impact of Unemployment via Joint Significance:

1) F-test: percenttotalunioncov and unemploymentrate

Restricted Model: (Refer to **Figure 11**in the **Appendix** for the STATA output)

population: meanhourwage = $\beta_0 + \beta_2$log_realgdppercapita + $\beta_4$gradratehs + $\beta_5$percentmanufacturing + $\beta_6$percentitprofessionalservices + $\beta_7$percentgraddegree + u sample: meanhourwage = $-123.0001 + 13.7090 \textit{log\_realgdppercapita} - 0.0370 \textit{gradratehs} - 17.4750 \textit{percentmanufacturing} + 46.9213 \textit{percentitprofessionalservices} + 0.0639 \textit{percentgraddegree}$

H_0: $\beta_1 = \beta_3 = 0$ **H_a = H_0 not true**

Unrestricted SSR = 79.5085477 Restricted SSR = 163.9217

df = 44 q = 2

5% critical value = 3.21 1% critical value = 5.12

F-stat = [(163.9217 – 79.5085477) / 79.5085477] * (44 / 2) = **23.357**

We reject the null hypothesis at both the 5% and 1% significant levels, which indicates that the variables for the percentage of total union coverage and the unemployment rate are jointly significant.

2) **F-test: log_realgdppercapita and unemploymentrate**

Restricted Model: (Refer to **Figure 12** in the **Appendix** for the STATA output)

population: meanhourwage = $\beta_0 + \beta_1$percenttotalunioncov + $\beta_4$gradratehs + $\beta_5$percentmanufacturing + $\beta_6$percentitprofessionalservices + $\beta_7$percentgraddegree + u

sample: meanhourwage = $11.1175 + 0.3101 \textit{percenttotalunioncov} + 0.0699 \textit{gradratehs} - 11.7421 \textit{percentmanufacturing} + 64.2038 \textit{percentitprofessionalservices} + 0.0178 \textit{percentgraddegree}$

H_0: $\beta_2 = \beta_3 = 0$ **H_a = H_0 not true**

Unrestricted SSR = 79.5085477 Restricted SSR = 106.308847

df = 44 q = 2

5% critical value = 3.21 1% critical value = 5.12

F-stat = [(106.308847 – 79.5085477) / 79.5085477] * (44 / 2) = **7.4156**

We reject the null hypothesis at both the 5% and 1% significant levels, which indicates that the variables for the log of the real GDP per capita and the unemployment rate are jointly significant.

Dummy Variables

We ran our regression including two dummy variables to capture the effect of variations in the minimum wage and in the Right to Work (RTW) policy on mean hourly wages. Both minimum wage and RTW are state mandated policies. We predict that states with a minimum wage above the federal minimum might have higher average hourly wages than states with a minimum wage lower than the federal wage or no minimum wage at all. RTW is a law that prohibits compulsory union membership—that is, employers and labor unions cannot require workers to pay union dues as a condition of employment. There is no consensus in the literature on how RTW affects wages due to the highly politicized nature of this law; however, it is often argued that states with RTW will experience lower wages compared to non-RTW states. This is because RTW reduces the strength of unions and thus might reduce their ability to put upward pressure on wages. We sought to test this in our model.

Model with Dummies: (Refer to **Figure 13** in the **Appendix** for the STATA output)

population: meanhourwage = $\beta_0 + \beta_1$percenttotalunioncov + $\beta_2$log_realgdppercapita + $\beta_3$unemploymentrate + $\beta_4$gradratehs + $\beta_5$percentmanufacturing + $\beta_6$percentitprofessionalservices + $\beta_7$percentgraddegree + $\beta_8$minwagehigh + $\beta_9$right2work + u

sample: meanhourwage = –83.0748 + 0.2126*percenttotalunioncov* + 9.2320*log_realgdppercapita* + 0.2797*unemploymentrate* + 0.0234*gradratehs* – 9.3678*percentmanufacturing* + 46.4404*percentitprofessionalservices* + 0.0297*percentgraddegree* – 0.0719*minwagehigh* – 0.86070*right2work*

1) Minimum wage, *coded as minwage*

Minimum wage is defined as the minimum hourly wage an employer can pay an employee. Currently, the federal minimum wage is set at $7.25. We have included minimum wage by state as a dummy variable to account for the differences within our observations for mean hourly wage. Minimum wage takes on a value of 0 if it is below the federal standard or if the state does not have a minimum wage. It takes a value of 1 if it is equal to or above the federal standard. This data is from the 2018 Consolidated Minimum Wage Table from the U.S. Department of Labor.

minwagehigh = 0 if < $7.25 (below federal minimum wage)

minwagehigh = 1 if ≥ $7.25 (above federal minimum wage)

2) Right to work, *coded as right2work*

Right to work is a state policy that guarantees every person the freedom to work for a living without mandatory union membership as a condition of employment. We include this as a dummy variable in our regression which takes on the value of 0 if a state is not a right to work state and a value of 1 if the state is a right to work state. This data is from the National Right to Work Legal Defense and Education foundation.

right2work = 0 if workers are required to be in a union

right2work = 1 if workers are NOT required to be in a union

Table 6. Dummy Variable Summary Statistics.

Dummy Variable Summary Statistics					
Variable	**Observation**	**Mean**	**St.dev**	**Min**	**Max**
minwagehigh	50	0.86	0.35	0	1
right2work	50	0.56	0.50	0	1

Final MLR Model:

For the final MLR model, we re-incorporated the variables that were jointly significant.

Table 7. Final MLR Model Results.

OLS Coefficient Results — Dependent Variable: meanhourwage	
Independent Variables	**Final MLR Model** df = 43 (c.v. 10% = 1.68) (c.v. 5% = 2.01) (c.v. 1% = 2.69)
percenttotalunioncov	0.2740 *** (6.66)
log_realgdppercapita	9.8985 *** (3.78)
unemploymentrate	0.3411 (1.25)
gradratehs	–0.0033 (–0.07)
percentitprofessionalservices	51.8871 *** (4.75)
percentgraddegree	0.0983 (1.27)
intercept	–91.3271 (–3.33)

*** significant at 1%, ** significant at 5%, * significant at 10%
NOTE: values in the parenthesis are t-statistics

Results

Being a state with a minimum wage above the federal minimum has no statistically significant impact on mean hourly wages relative to states with a minimum wage below the federal minimum. This could be due to the fact that differences in minimum wage are accounted for within mean hourly wage rates by state. Additionally, after looking at a scatter plot between our coded dummy variable and mean hourly wage, we do not see any noticeable relationship between a minimum wage above the federal standard and a higher mean hourly wage, which could also be a cause of insignificant results. (Refer to **Figure 14** in the **Appendix**)

A right to work state did not have a statistically significant impact on hourly wages relative to states without right to work. As previously mentioned, this policy is highly politicized and has continued to change over the past decade as more states become right to work states by statute or constitutional provisions. The evolving nature of this policy could mean that not enough time has passed to be able to conclude with any definitive effects this policy has had.

After running all of our robustness checks and dummy variables, our final model is as follows:
(Refer to **Figure 15** in the **Appendix** for the STATA output)

population: meanhourwage = $\beta_0 + \beta_1$percenttotalunioncov + $\beta_2$log_realgdppercapita + $\beta_3$unemploymentrate + $\beta_4$gradratehs + $\beta_5$percentitprofessionalservices + $\beta_6$percentgraddegree + u

sample: meanhourwage = –91.3271 + 0.2740*percenttotalunioncov* + 9.8985*log_realgdppercapita* + 0.3411*unemploymentrate* – 0.0033*gradratehs* + 51.8871*percentitprofessionalservices* + 0.0983*percentgraddegree*

VI. Conclusions

Our goal for conducting this research was to identify the effects of unionization on hourly wages across all industries at the state level. Of the seven explanatory variables in our Full MLR Model, three variables, including unionization, showed statistically significant results. The percentage of the economy captured by the IT sector had the largest impact on wages, followed by unionization, and finally GDP. High school graduation rate, percentage of workers with a master's degree or higher, and unemployment were not independently significant but were jointly significant with one of the three variables mentioned above.

The notable size of the beta coefficient of the percentage of employees in the IT sector implies that wages are heavily influenced by high paying sectors. This result stresses the need for further analysis and the creation of an industry index. An industry index would allow us to account for the composition of high- and low-paying jobs in the state in order to decrease the omitted variable bias that is likely caused from leaving out other high paying industries similar to IT. Although manufacturing did not show a significant effect on wages, this does not negate the need to account for the percentage of low-paying sectors in the economy. Manufacturing might not have been the best variable to use as a proxy for low-paying jobs. Alternatively, we could improve upon our dataset by finding industry-specific wage data, eliminating the need for an industry index all together and increasing our sample size.

It is interesting to note that variables measuring education, on their own, were insignificant in all models. However, multiple joint significance tests measuring the impact of education demonstrated statistically significant results, implying that education does have an impact on wages. This likely occurred because of multicollinearity.

With the inclusion of dummy variables in our model, we sought to explore the effect that minimum wage rates above and below the federal standard would have on wages. Although one would intuitively believe that minimum wage rates would affect hourly wages, results from our regression model show that it is statistically insignificant. Furthermore, plotting the dummy variable against mean hourly wages shows that although this logic holds in the case of states with minimum wages below the federal standard (i.e., in general, states with lower minimum wages appear to have lower mean hourly wages), it does not hold for states with a federal minimum wage above the standard (i.e., the distribution for minimum wages above the federal wage appears more random).

Going forward, beyond including an industry index, we suggest exploring the effect of demographic factors on wages in the model, such as gender and race. Unions tend to over-represent certain demographics—males being one example—and under-represent others. It might be relevant to create different dummy variables for demographics to use as an interaction term with unionization.

VII. References

Blanchflower, David and Alex Bryson (2007). “The Wage Impact of Trade Unions in the UK Public and Private Sectors.” *IZA Discussion Paper No. 3055.*

Card, D. (2001). The Effect of Unions on Wage Inequality in the U.S. Labor Market. *Industrial and Labor Relations Review, 54*(2), 296-315. doi:10.2307/2696012.

Farber, H. (2001). Notes on the Economics of Labor Unions. *Princeton University Industrial Relations Section Working Paper #452.*

Freeman, Richard B., and James L. Medoff (1985). What Do Unions Do? *Industrial and Labor Relations Review*, Vol. 38 No. 2, Cornell University. https://journals.sagepub.com/doi/abs/10.1177/001979398503800207.

Freeman, Richard Barry, and Morris M. Kleiner (1990). The impact of new unionization on wages and working conditions. *Journal of Labor Economics,* 8(1) Part 2: S8-S25.

Oswald, Andrew (1982). “Trade Unions, Wages and Unemployment: What Can Simple Models Tell Us.” *Oxford Economic Papers,* New Series, Vol. 34, No. 3.

Pencavel, John H., and Catherine E. Hartsog (1984). A Reconsideration of the Effects of Unionism on Relative Wages and Employment in the United States, 1920-80. *Journal of Labor Economics*, Vol. 2, No. 2, pp. 193-232.

Rosen, Sherwin (1994). "H. Gregg Lewis and Modern Labor Economics," *Journal of Labor Economics*, University of Chicago Press, vol. 12(1), pages 139-143, January.

Western, Bruce and Jake Rosenfeld (2011) Unions, Norms, and Rise in U.S. Wage Inequality. *American Sociological Review*, Vol 76: Iss 4, pp. 513-537. doi: https://doi.org/10.1177/0003122411414817.

VIII. Appendix

Figure 1. Scatterplot of our dependent variable and our main variable of interest (meanhourlywage vs. percenttotalunioncov).

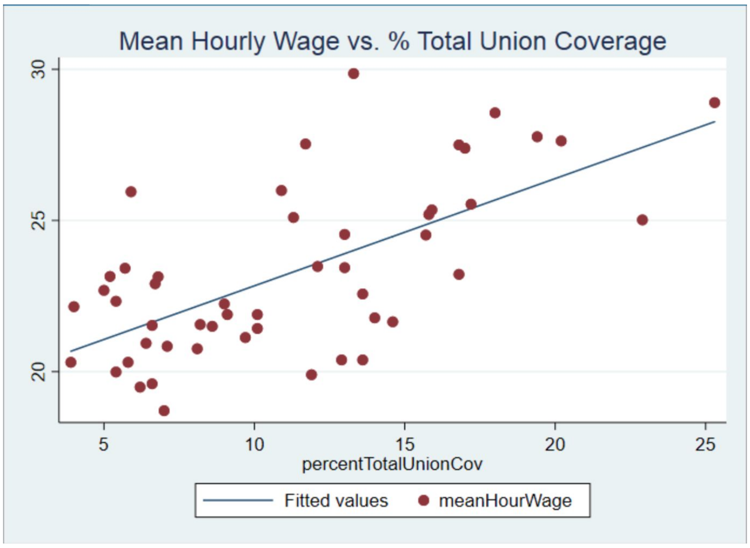
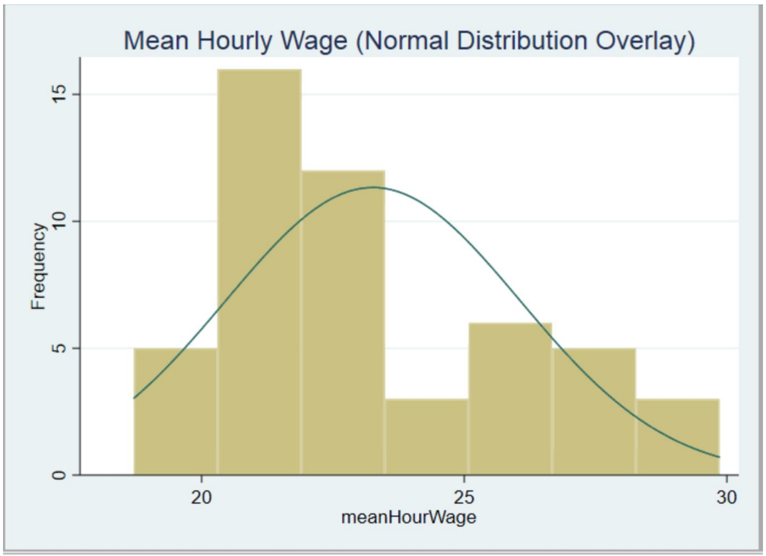


Figure 2. Distribution of the Mean Hourly Wage Variable (across states, after dropping D.C.).



RESIDUAL PLOTS

Figure 3. SLR Model Residual Plot and Full MLR Model Residual Plot.

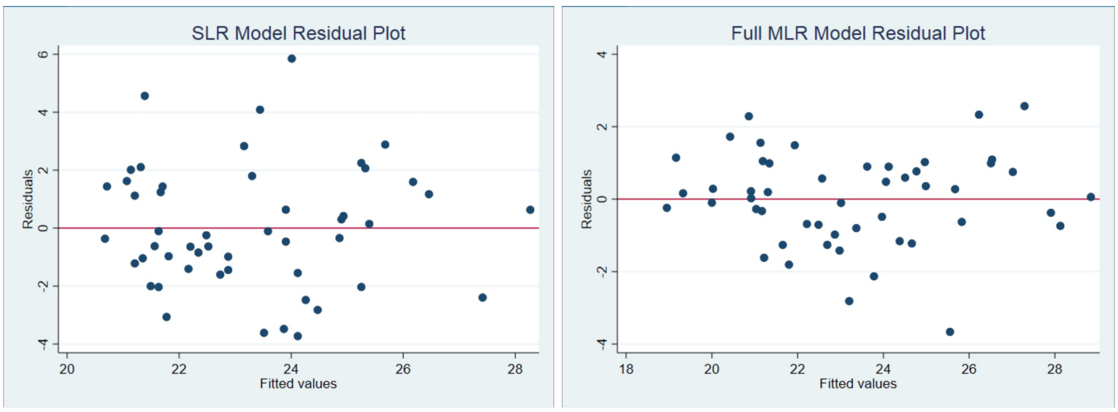


Figure 4. Modified MLR Model Residual Plot and Final MLR Model Residual Plot.

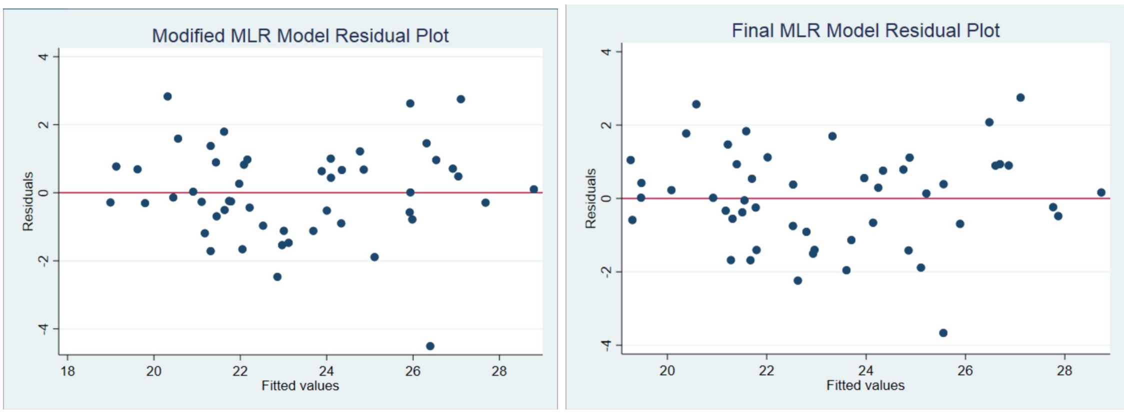


Figure 5. Stata Output – Simple Linear Regression Model.

```
. regress meanhourwage percenttotalunioncov

      Source |       SS           df       MS      Number of obs   =        50
-------------+----------------------------------   F(1, 48)        =     35.94
       Model |  164.918753         1  164.918753   Prob > F        =    0.0000
    Residual |   220.23256        48  4.58817833   R-squared       =    0.4282
-------------+----------------------------------   Adj R-squared   =    0.4163
       Total |  385.151312        49  7.86023086   Root MSE        =     2.142

--------------------------------------------------------------------------------------
        meanhourwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
---------------------+----------------------------------------------------------------
percenttotalunioncov |   .3547526   .0591713     6.00   0.000     .2357809    .4737244
               _cons |   19.29192    .728131    26.50   0.000     17.82791    20.75592
--------------------------------------------------------------------------------------
```

Figure 6. Stata Output – Multiple Linear Regression Model.

```
. regress meanhourwage percenttotalunioncov log_realgdppercapita unemploymentrate gradratehs percentmanufacturing
> percentitprofessionalservices percentgraddegree

      Source |       SS           df       MS      Number of obs   =        50
-------------+----------------------------------   F(7, 42)        =     23.06
       Model |  305.642765         7  43.6632521   Prob > F        =    0.0000
    Residual |  79.5085477        42  1.89306066   R-squared       =    0.7936
-------------+----------------------------------   Adj R-squared   =    0.7592
       Total |  385.151312        49  7.86023086   Root MSE        =    1.3759

-----------------------------------------------------------------------------------------------
                 meanhourwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
------------------------------+----------------------------------------------------------------
         percenttotalunioncov |   .2652871   .0409212     6.48   0.000     .1827048    .3478694
         log_realgdppercapita |   9.683926   2.580403     3.75   0.001     4.476461    14.89139
             unemploymentrate |   .2978092   .2692288     1.11   0.275    -.2455165    .8411349
                   gradratehs |   .0216331   .0508932     0.43   0.673    -.0810736    .1243398
         percentmanufacturing |  -10.21135    6.66902    -1.53   0.133    -23.66998    3.247279
percentitprofessionalservices |   48.62262   10.96466     4.43   0.000     26.49504    70.75021
            percentgraddegree |   .0669063   .0786539     0.85   0.400    -.0918236    .2256361
                        _cons |  -89.44989   27.01662    -3.31   0.002    -143.9716   -34.92815
-----------------------------------------------------------------------------------------------
```

Figure 7. Stata Output – Modified Multiple Linear Regression Model.

```
. regress meanhourwage percenttotalunioncov log_realgdppercapita percentitprofessionalservices

      Source |       SS           df       MS      Number of obs   =        50
-------------+----------------------------------   F(3, 46)        =     50.30
       Model |  295.167847         3  98.3892824   Prob > F        =    0.0000
    Residual |   89.983465        46  1.95616228   R-squared       =    0.7664
-------------+----------------------------------   Adj R-squared   =    0.7511
       Total |  385.151312        49  7.86023086   Root MSE        =    1.3986

-----------------------------------------------------------------------------------------------
                 meanhourwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
------------------------------+----------------------------------------------------------------
         percenttotalunioncov |   .2852374   .0400332     7.13   0.000     .2046546    .3658201
         log_realgdppercapita |   8.383756   2.364196     3.55   0.001     3.624876    13.14264
percentitprofessionalservices |   58.63643   10.07317     5.82   0.000      38.3602    78.91267
                        _cons |  -73.45204   25.16831    -2.92   0.005    -124.1132   -22.79086
-----------------------------------------------------------------------------------------------
```

F-TEST OUTPUTS

Figure 8. Stata Output – Education F-Test – Restricted Model 1: percentage of graduate degrees and percentage of IT sector

```
. regress meanhourwage percenttotalunioncov log_realgdppercapita unemploymentrate gradratehs percentmanufacturi
> ng

      Source |       SS           df       MS      Number of obs   =        50
-------------+----------------------------------   F(5, 44)        =     18.25
       Model |  259.863564         5  51.9727128   Prob > F        =    0.0000
    Residual |  125.287748        44  2.84744882   R-squared       =    0.6747
-------------+----------------------------------   Adj R-squared   =    0.6377
       Total |  385.151312        49  7.86023086   Root MSE        =    1.6874

--------------------------------------------------------------------------------------
        meanhourwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
---------------------+----------------------------------------------------------------
percenttotalunioncov |   .2571263   .0501184     5.13   0.000     .1561193    .3581333
log_realgdppercapita |   13.66327   2.925569     4.67   0.000     7.767176    19.55937
    unemploymentrate |   .3737124   .3263339     1.15   0.258    -.2839703    1.031395
          gradratehs |    .015188   .0623317     0.24   0.809    -.1104334    .1408093
percentmanufacturing |  -20.56927   7.569758    -2.72   0.009    -35.82512   -5.313428
               _cons |  -127.5642   30.67071    -4.16   0.000    -189.3769   -65.75141
--------------------------------------------------------------------------------------
```

Figure 9. Stata Output – Education F-Test – Restricted Model 2: high school graduation rate, percentage of graduate degrees, and GDP

```
. regress meanhourwage percenttotalunioncov unemploymentrate percentmanufacturing percentitprofessionalservices
>

      Source |       SS           df       MS      Number of obs   =        50
-------------+----------------------------------   F(4, 45)        =     27.92
       Model |  274.542114         4  68.6355286   Prob > F        =    0.0000
    Residual |  110.609198        45  2.45798217   R-squared       =    0.7128
-------------+----------------------------------   Adj R-squared   =    0.6873
       Total |  385.151312        49  7.86023086   Root MSE        =    1.5678

-----------------------------------------------------------------------------------------------
                 meanhourwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
------------------------------+----------------------------------------------------------------
         percenttotalunioncov |   .3071252   .0445601     6.89   0.000     .2173764    .3968739
             unemploymentrate |  -.0558045   .2833626    -0.20   0.845    -.6265262    .5149172
         percentmanufacturing |  -8.797829   6.925183    -1.27   0.210    -22.74586    5.150206
percentitprofessionalservices |   65.73743   11.23942     5.85   0.000     43.10009    88.37478
                        _cons |   17.09854   1.645509    10.39   0.000     13.78431    20.41276
-----------------------------------------------------------------------------------------------
```

Figure 10. Stata Output – Education F-Test – Restricted Model 3: high school graduation rate and GDP

```
. regress meanhourwage percenttotalunioncov unemploymentrate percentmanufacturing percentitprofessionalservices
> percentgraddegree

      Source |       SS           df       MS      Number of obs   =        50
-------------+----------------------------------   F(5, 44)        =     21.88
       Model |  274.660838         5  54.9321675   Prob > F        =    0.0000
    Residual |  110.490475        44  2.51114715   R-squared       =    0.7131
-------------+----------------------------------   Adj R-squared   =    0.6805
       Total |  385.151312        49  7.86023086   Root MSE        =    1.5847

------------------------------------------------------------------------------------------------
                 meanhourwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
------------------------------+-----------------------------------------------------------------
         percenttotalunioncov |   .3073388   .0450502     6.82   0.000     .2165461    .3981315
             unemploymentrate |  -.0494937   .2878776    -0.17   0.864    -.6296728    .5306854
         percentmanufacturing |   -8.36991   7.271079    -1.15   0.256    -23.02381    6.283987
percentitprofessionalservices |   65.15477   11.67209     5.58   0.000     41.63121    88.67832
            percentgraddegree |   .0194451   .0894292     0.22   0.829    -.1607875    .1996778
                        _cons |   16.85391   2.007994     8.39   0.000     12.80706    20.90075
------------------------------------------------------------------------------------------------
```

Figure 11. Stata Output – Unemployment F-Test – Restricted Model 1: unemployment and union coverage

```
. regress meanhourwage log_realgdppercapita gradratehs percentmanufacturing percentitprofessionalservices perce
> ntgraddegree

      Source |       SS           df       MS      Number of obs   =        50
-------------+----------------------------------   F(5, 44)        =     11.88
       Model |  221.229612         5  44.2459224   Prob > F        =    0.0000
    Residual |    163.9217        44  3.72549319   R-squared       =    0.5744
-------------+----------------------------------   Adj R-squared   =    0.5260
       Total |  385.151312        49  7.86023086   Root MSE        =    1.9302

------------------------------------------------------------------------------------------------
                 meanhourwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
------------------------------+-----------------------------------------------------------------
         log_realgdppercapita |   13.70904   3.384513     4.05   0.000     6.888002    20.53008
                   gradratehs |  -.0369973   .0688982    -0.54   0.594    -.1758524    .1018579
         percentmanufacturing |  -17.47503   9.200654    -1.90   0.064    -36.01773    1.067673
percentitprofessionalservices |   46.92127   15.26965     3.07   0.004      16.1473    77.69524
            percentgraddegree |   .0639151   .1094003     0.58   0.562    -.1565668     .284397
                        _cons |  -123.0001   34.91741    -3.52   0.001    -193.3716   -52.62874
------------------------------------------------------------------------------------------------
```

Figure 12. Stata Output – Unemployment F-Test – Restricted Model 2: unemployment and log_gdp

```
. regress meanhourwage percenttotalunioncov gradratehs percentmanufacturing percentitprofessionalservices perce
> ntgraddegree

      Source |       SS           df       MS      Number of obs   =        50
-------------+----------------------------------   F(5, 44)        =     23.08
       Model |  278.842466         5  55.7684931   Prob > F        =    0.0000
    Residual |  106.308847        44  2.41611015   R-squared       =    0.7240
-------------+----------------------------------   Adj R-squared   =    0.6926
       Total |  385.151312        49  7.86023086   Root MSE        =    1.5544

------------------------------------------------------------------------------------------------
                 meanhourwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
------------------------------+-----------------------------------------------------------------
         percenttotalunioncov |   .3100606   .0442297     7.01   0.000     .2209215    .3991997
                   gradratehs |   .0699212   .0526834     1.33   0.191    -.0362553    .1760977
         percentmanufacturing |  -11.74213   7.488835    -1.57   0.124    -26.83488    3.350629
percentitprofessionalservices |   64.20383   11.46971     5.60   0.000     41.08814    87.31952
            percentgraddegree |   .0178028   .0873068     0.20   0.839    -.1581524     .193758
                        _cons |   11.11752   4.421611     2.51   0.016     2.206348    20.02869
------------------------------------------------------------------------------------------------
```

Figure 13. Stata Output – Full MLR Model with Dummy Variables

```
. regress meanhourwage percenttotalunioncov log_realgdppercapita unemploymentrate gradratehs percentmanufacturi
> ng percentitprofessionalservices percentgraddegree minwagehigh right2work

      Source |       SS           df       MS      Number of obs   =        50
-------------+----------------------------------   F(9, 40)        =     17.97
       Model |   308.79071         9  34.3100789   Prob > F        =    0.0000
    Residual |  76.3606024        40  1.90901506   R-squared       =    0.8017
-------------+----------------------------------   Adj R-squared   =    0.7571
       Total |  385.151312        49  7.86023086   Root MSE        =    1.3817

------------------------------------------------------------------------------------------------
                  meanhourwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------------------------+----------------------------------------------------------------
           percenttotalunioncov |   .2126241   .0623372     3.41   0.001     .0866359    .3386123
          log_realgdppercapita |   9.231976   2.622148     3.52   0.001     3.932418    14.53153
              unemploymentrate |   .2797176   .2786977     1.00   0.322    -.2835515    .8429867
                    gradratehs |   .0234118   .0517794     0.45   0.654    -.0812382    .1280619
           percentmanufacturing |  -9.367862   6.739374    -1.39   0.172    -22.98865    4.252921
percentitprofessionalservices |   46.44042   11.32483     4.10   0.000     23.55209    69.32875
             percentgraddegree |   .0297281   .0848834     0.35   0.728    -.1418276    .2012838
                   minwagehigh |  -.0719022   .6682762    -0.11   0.915    -1.422539    1.278734
                     right2work |   -.860702   .6725837    -1.28   0.208    -2.220044    .4986404
                          _cons |  -83.0748   27.63385    -3.01   0.005    -138.9249   -27.22471
------------------------------------------------------------------------------------------------
```

Figure 14. Scatter plot of minwagehigh dummy variable against meanhourwage

Figure 15. Stata Output – Final Multiple Linear Regression Model.

```
. regress meanhourwage percenttotalunioncov log_realgdppercapita unemploymentrate gradratehs percentitprofessio
> nalservices percentgraddegree

      Source |       SS           df       MS      Number of obs   =        50
-------------+----------------------------------   F(6, 43)        =     25.71
       Model |  301.204567         6  50.2007612   Prob > F        =    0.0000
    Residual |  83.9467453        43  1.95224989   R-squared       =    0.7820
-------------+----------------------------------   Adj R-squared   =    0.7516
       Total |  385.151312        49  7.86023086   Root MSE        =    1.3972

------------------------------------------------------------------------------------------------
                  meanhourwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------------------------+----------------------------------------------------------------
           percenttotalunioncov |   .2740216   .0411502     6.66   0.000     .1910342     .357009
          log_realgdppercapita |   9.898489   2.616566     3.78   0.000     4.621681     15.1753
              unemploymentrate |   .3410756   .2718953     1.25   0.216    -.2072535    .8894046
                    gradratehs |  -.0033361   .0489578    -0.07   0.946    -.1020688    .0953966
percentitprofessionalservices |   51.88706   10.92224     4.75   0.000     29.86026    73.91385
             percentgraddegree |    .098284   .0771153     1.27   0.209    -.0572339    .2538018
                          _cons |  -91.32708   27.40746    -3.33   0.002    -146.5995   -36.05467
------------------------------------------------------------------------------------------------
```

SCATTER PLOTS OF MODEL CHECKS

Figure 16. SLR Model Check: Plot of Observed vs. Predicted and Full MLR Model Check: Plot of Observed vs. Predicted.

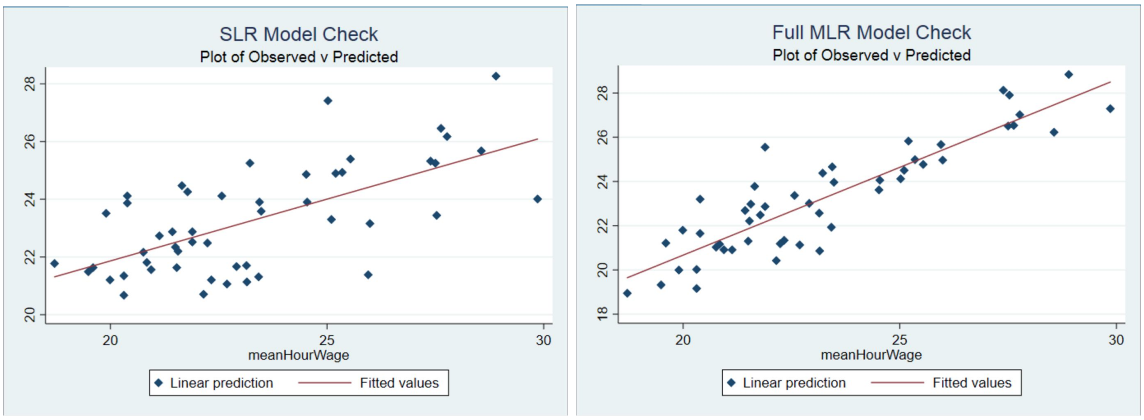


Figure 17. Modified MLR Model Check: Plot of Observed vs. Predicted and Final MLR Model Check: Plot of Observed vs. Predicted.