Right-To-Work: An Empirical Analysis and Comparison Of The Wage Differentials In States That Have Adopted Right-To-Work Laws and Those Who Have Not

Trevor Carnes

5/3/2020

## Summary

Labor unions, and their effects on wages in a state has been the subject of debate since their early beginnings in the 1930s. In this research project, I begin with this simple question: are the average wages higher in states that have adopted RTW (Right-To-Work) laws, or are they higher in states that have not adopted RTW laws. Subsequently I propose model H\_o: RTWStateWagesNonRTWStateWages, In my attempt to answer this question, I collect data from various sources such as the Bureau of Labor Statistics, the Current Populations Survey, The Job Openings and Labor Turnover Survey and more. I have collected wage, employment and education data for years dating back to roughly 2009. To answer these questions it is important to understand what unions are, their intended purpose, and what RTW laws are and their intended purpose and the history behind the matters  
 In 1935 congress enacted the National Labor Relations Act in an effort to quell the growing inequality between labors bargaining power, and the labor class share of income, that was being further exacerbated by the Great Depression. This law guarantees the rights of employees who wish to form a union and participate in collective bargaining procedures, that is the workers ability to negotiate wages, benefits and health and safety protections for the unionized workers.  
It wasn’t until 12 years later that congress passed the Labor Management Relations act of 1947, overriding president Harry Truman’s veto. This act repealed some aspects of the NLRA (National Labor Relations Act), the main one being the outlawing of “closed shops”. The premise of a closed shop is that employment was contingent on joining a union. This act would soon pave the way for what would become Right-To-Work laws, which would be voted on and administered on a state-by-state basis. Like the LMRA (Labor Management Relations Act) RTW laws go a step further and ban what has become known as agency shop, which requires employees to pay union representation costs whether they are members of that union or not. Wisconsin, West Virginia, Indiana and Kentucky are among the states that have recently adopted these right-to-work laws. Other states such as Arizona, Arkansas and Alabama have had RTW laws on the books since the late 1940s, shortly aster the passing of the LMRA.

## Literature Review

 Though this topic is hotly debated, the existing literature of RTW laws on state wages level is fairly limited. In a more recent 2010 paper, Steven Popejoy of the University of Central Missouri purposed that union wages were determined by the supply and demand for union labor. He modeled a supply and demand function for union labor based off the following model. Let U\_d be the demand for union labor services and U\_s be the supply of union services.

U\_d=U\_d (p,w,d,np,t) & U\_s=U\_s (p,c,g) Where p = price of union membership w = wage rate as a proxy for wealth d = union vs. non-union wage differential np = non-pecuniary benefits of union members t = tastes and preferences  
c = cost of organizing and providing services g = goals of the union | Dissecting the proposed function, we would expect the price of union membership to hold a negative correlation to demand for services. Price of union membership may take the form of union dues and fees that a union may require to cover the cost of collective bargaining. The wage rate as a proxy for wealth would be expected to carry a positive coefficient, as it rises the demand for union services may rise as well because the union will be offer higher compensation for membership. The union wage differential is the difference in wages between those in a union and those not in a union, this may be assumed to have a positive correlation with the demand for services because the higher that difference the more attractive union services look to those not in a union thus the demand will grow. Non-pecuniary benefits would be positively correlated since the better the benefits the more people will want to join the union. Also taste and preferences, cost of organizing and goals of the union maybe subjective in their correlation and may only be determined from empirical collected from with in the states. Popejoy found that total unionism in a state would be found when union demand and union supply are found in equilibrium such that, U=U\_s=U\_D. | Net migration may play a crucial role in the supply and demand of union labor and the determination of wages. In a paper published in the Cato Journal Richard Vedder of Ohio State University makes the claim that there has been a large net migration of people living in non-RTW states too RTW states. He cites that in 1970 28.5 percent of the U.S. population lived in RTW states. However, from U.S. census bureau data, in Vedder found that the percentage living in RTW states had increased to almost 40 percent in 2008, that’s nearly 121 million people. | Vedder also makes the claim that most important reason for the increase in the percentage of U.S. population living in right-to-work states is because there has been a large net migration from non-RTW states to those states who have laws that allow for a persons discretion in regard to joining a union. A direct correlation between RTW and net migration be may not be that obvious, but the assertion that economic growth in these states are high may be much more likely explanation. Therefore internal in-migration has grown greatly. Vedder states that 4.7 million Americans migrated from non-RTW states to RTW states from April 1st 2000 to July 1st 2008. He cites that reduces down to be more than one person every minute to these states. | In his paper after summing up the effects of RTW on net migration he then goes on to make the economic case for the advantages of RTW laws. Vedder asserts that RTW laws can act as a catalyst to significant economies of scale for labor and employment in these states. His evidence suggests that RTW laws increase the availability of factor endowments that lead to greater employment and labor productivity, particularly the latter. According to Vedder, labor is the single most import part of production. The owners of labor reap the benefits of their production at nearly twice the rate then the owners of the other factor inputs. For further analysis of the importance of labor as a factor input, Vedder employees the Cobb-Douglas Production Function. Using this function, he estimates that the elasticity of output with respect labor is about .7. | This suggests that a 1 percent increase in labor as an input will lead a .7 increase in output. So as a result, in states with RTW laws, labor inputs increase roughly 2 to 3 percent. Therefore according to his estimations labor output as a whole should benefit to the tune of 2 percent. In summation of Vedders work and calculations, he states that the above cited number came from a sample of 40 states. If the positive externalities of RTW laws were spread to all 50 states, the aggregate national output would be roughly .8 percent, roughly 110 billion dollars a year, which would equal an over 1000 dollar increase for a typical house hold of 3. | Vedder then uses multiple regression analysis to relate the real growth in per-capita personal income from 1977 to 2007 to whether or not RTW laws had been enacted within the states. Aside from the RTW variable he includes the following for variables for control purposes. The first two are tax variables. The first is TAXBURDEN77 to account for the amount of state and local taxes paid within the state. Next was CHTAXBURDEN, which accounts for the change in the above tax burden from 1977 to 2007. He then included a COLLEGE variable which gives the proportion of the population which has attained a 4 year college degree or higher. He also included a LANDAREA variable that takes into account the amount of land area as a factor input a state may have. Lastly, Vedder includes a population POPGROWTH variable to control for any large growths in population for the regressed state.

 From his regression results, suppose that two sates were identical aside from one having RTW laws and the other one not having these laws. Along with this Vedder says to assume that the non-RTW have real personal income growth or roughly 50 percent from 1977 to 2007. His model would shows that the RTW state had about 23 percent higher growth rate than its non-RTW counterparts. Along with the previous assumptions, suppose the previously regressed states at real income per-capita income of 24,000 in 2007 dollars. The income in the non-RTW state would have grown to roughly 36,000 in contrast with states that adopted RTW laws who’s per-capita incomes would have grown roughly 38,760 in 2007 dollars. According to his model, the RTW laws increased per-capita income by 2,760, or when divided among the population of the state grouped in families of 3, real income for families rose almost 11,000 dollars holding all else constant. Finally Vedder points to the RTW passage and its effect in Louisiana and Idaho as hard evidence of his conclusion.  
 In contrast to Vedders finding that RTW laws have positive effects on labor and employment, Larry Mishel of the Economic Policy institute found that RTW laws have harmful effects on employment and wages within a state. He estimated a log wage equation based of data from the Bureau of labor Statistics current population survey. His sample consisted of 152,576 prime age workers (that are workers between the ages of 18 & 64) for the year 2000. The average hourly wage for those workers was 15.54 dollars an hour and the median 12.24 dollars an hour. Mishel found that the average wage in the RTW states were 11.45 dollars an hour and 13 dollars an hour for non-RTW states. This is nearly a 11.9% wage deficit for workers in RTW states. Since the literature is mixed, contrary to popular belief, the need for further research into this matter, particularly at the state wage level, is needed.

## Data and Methodology

 To estimate wages in RTW states and Non-RTW states, I use multiple regression analysis, in which I propose multiple models to estimate mean state wages and annual salaries. The variables I obtained and the correlations I hypothesize they will carry are listed below. For each variable, I have over 500 observations of all 50 states dating back for nearly 10 years.

• Hourly Mean: The mean hourly wage in a state in a given year. I use this a dependent variable on which I regress the independent variables.  
• Hourly Median: The median wage within state. This may also be used an independent variable and used for robustness testing. • Annual Mean: This measures the mean annual salary within a state in a given year, also used as a dependent variable. • Annual Median: A measure for the median salary with a state in a given year. As with the hourly median, this can be used as a dependent variable or used only for robustness testing. • Total Employment: Measures the number of total employed people within a state. I would expect this variable to carry a negative correlation with mean hourly wages and mean annual salaries. As total employment increases, the demand for labor will fall, causing a decrease in wages and salaries. • Job Openings: The number of open jobs in a state in a given year. I expect this to carry a positive coefficient since this may represent demand for labor. The higher the demand for labor, the higher the wages being offered. • Hires: The number of hires within a state. The correlation with this will be one of interest. • Quits: The number of quits within a state in a given year. I expect this to carry a negative correlation, since the more people who quit inundates the supply of labor and lowers wages. • Layoffs and Discharges: The number of people who were either laid off or discharged. This one should follow suit with the quits variable. • Total Separations: This accounts for all fires, quits, layoffs and discharges within a state. • Children: Percentage of the population who are children. • No High: Percentage of the population that have not achieved a high school education. I expect this variable to cause a drag effect on wages and be negatively correlated.  
• High: Percentage of the population that have a high school diploma as their highest form of education. Since education leads to higher wages, I would expect this one to be positively correlated but with a low coefficient • Some: Percentage of the population that have some college. I expect this one to be positively correlated but not much higher coefficient than high school.  
• Bachelors: The percentage of the population that has a bachelor’s degree s their highest form education. This one I expect to carry a strong positive correlation. • RTW: A binary variable, where 1 means the state has right to work, laws and 2 means it does not. I expect this to carry a negative coefficient. I collected all this data from various sources such as Bureau of Labor Statistics, Job Openings and Labor Turn Over Survey, and Current population survey. These are the packages installed and used to create this report

library(jtools)  
library(rio)  
library(usmap)  
library(rmarkdown)  
library(shiny)  
library(stargazer)

##   
## Please cite as:

## Hlavac, Marek (2018). stargazer: Well-Formatted Regression and Summary Statistics Tables.

## R package version 5.2.2. https://CRAN.R-project.org/package=stargazer

library(panelr)

## Loading required package: lme4

## Loading required package: Matrix

##   
## Attaching package: 'lme4'

## The following object is masked from 'package:rio':  
##   
## factorize

##   
## Attaching package: 'panelr'

## The following object is masked from 'package:stats':  
##   
## filter

library(tidyverse)

## Registered S3 methods overwritten by 'broom':  
## method from   
## tidy.glht jtools  
## tidy.summary.glht jtools

## ── Attaching packages ───────────────────────────────────────────────────────────────── tidyverse 1.3.0 ──

## ✓ ggplot2 3.3.0 ✓ purrr 0.3.3  
## ✓ tibble 2.1.3 ✓ dplyr 0.8.5  
## ✓ tidyr 1.0.2 ✓ stringr 1.4.0  
## ✓ readr 1.3.1 ✓ forcats 0.4.0

## ── Conflicts ──────────────────────────────────────────────────────────────────── tidyverse\_conflicts() ──  
## x tidyr::expand() masks Matrix::expand()  
## x dplyr::filter() masks panelr::filter(), stats::filter()  
## x dplyr::lag() masks stats::lag()  
## x tidyr::pack() masks Matrix::pack()  
## x tidyr::unpack() masks Matrix::unpack()

library(dplyr)  
library(readxl)  
library(ggplot2)

Below is the portion I collect from BLS with the documents I used and how I cleaned the data

WageData19 <- read\_excel("state\_M2019\_dl.xlsx" , na = "")  
WageData19 <- select(WageData19, area\_title, occ\_title, tot\_emp, h\_mean, a\_mean, h\_median, a\_median)   
WageData19 <- filter(WageData19, occ\_title == "All Occupations")  
WageData19 <- select(WageData19, - occ\_title)  
WageData19 <- mutate(WageData19, YEAR = 2019)  
WageData19 <- rename(WageData19, STATE = area\_title)  
  
WageData18 <- read\_excel("state\_M2018\_dl.xlsx" , na = "")  
WageData18 <- select(WageData18, STATE, OCC\_TITLE, TOT\_EMP, H\_MEAN, A\_MEAN, H\_MEDIAN, A\_MEDIAN)   
WageData18 <- filter(WageData18, OCC\_TITLE == "All Occupations")  
WageData18 <- select(WageData18, - OCC\_TITLE)  
WageData18 <- mutate(WageData18, YEAR = 2018)  
  
  
WageData17 <- read\_excel("state\_M2017\_dl.xlsx" , na = "")  
WageData17 <- select(WageData17, STATE, OCC\_TITLE, TOT\_EMP, H\_MEAN, A\_MEAN, H\_MEDIAN, A\_MEDIAN)   
WageData17 <- filter(WageData17, OCC\_TITLE == "All Occupations")  
WageData17 <- select(WageData17, - OCC\_TITLE)  
WageData17 <- mutate(WageData17, YEAR = 2017)  
  
  
WageData16 <- read\_excel("state\_M2016\_dl.xlsx" , na = "")  
WageData16 <- select(WageData16, STATE, OCC\_TITLE, TOT\_EMP, H\_MEAN, A\_MEAN, H\_MEDIAN, A\_MEDIAN)   
WageData16 <- filter(WageData16, OCC\_TITLE == "All Occupations")  
WageData16 <- select(WageData16, - OCC\_TITLE)  
WageData16 <- mutate(WageData16, YEAR = 2016)  
  
  
WageData15 <- read\_excel("state\_M2015\_dl.xlsx" , na = "")  
WageData15 <- select(WageData15, STATE, OCC\_TITLE, TOT\_EMP, H\_MEAN, A\_MEAN, H\_MEDIAN, A\_MEDIAN)   
WageData15 <- filter(WageData15, OCC\_TITLE == "All Occupations")  
WageData15 <- select(WageData15, - OCC\_TITLE)  
WageData15 <- mutate(WageData15, YEAR = 2015)  
  
  
WageData14 <- read\_excel("state\_M2014\_dl.xlsx" , na = "")  
WageData14 <- select(WageData14, STATE, OCC\_TITLE, TOT\_EMP, H\_MEAN, A\_MEAN, H\_MEDIAN, A\_MEDIAN)   
WageData14 <- filter(WageData14, OCC\_TITLE == "All Occupations")  
WageData14 <- select(WageData14, - OCC\_TITLE)  
WageData14 <- mutate(WageData14, YEAR = 2014)  
  
  
WageData13 <- read\_excel("state\_M2013\_dl.xls" , na = "")  
WageData13 <- select(WageData13, STATE, OCC\_TITLE, TOT\_EMP, H\_MEAN, A\_MEAN, H\_MEDIAN, A\_MEDIAN)   
WageData13 <- filter(WageData13, OCC\_TITLE == "All Occupations")  
WageData13 <- select(WageData13, - OCC\_TITLE)  
WageData13 <- mutate(WageData13, YEAR = 2013)  
  
  
WageData12 <- read\_excel("state\_M2012\_dl.xls" , na = "")  
WageData12 <- select(WageData12, STATE, OCC\_TITLE, TOT\_EMP, H\_MEAN, A\_MEAN, H\_MEDIAN, A\_MEDIAN)   
WageData12 <- filter(WageData12, OCC\_TITLE == "All Occupations")  
WageData12 <- select(WageData12, - OCC\_TITLE)  
WageData12 <- mutate(WageData12, YEAR = 2012)  
  
  
WageData11 <- read\_excel("state\_M2011\_dl.xls" , na = "")  
WageData11 <- select(WageData11, STATE, OCC\_TITLE, TOT\_EMP, H\_MEAN, A\_MEAN, H\_MEDIAN, A\_MEDIAN)   
WageData11 <- filter(WageData11, OCC\_TITLE == "All Occupations")  
WageData11 <- select(WageData11, - OCC\_TITLE)  
WageData11 <- mutate(WageData11, YEAR = 2011)  
  
  
WageData10 <- read\_excel("state\_M2010\_dl.xls" , na = "")  
WageData10 <- select(WageData10, STATE, OCC\_TITLE, TOT\_EMP, H\_MEAN, A\_MEAN, H\_MEDIAN, A\_MEDIAN)   
WageData10 <- filter(WageData10, OCC\_TITLE == "All Occupations")  
WageData10 <- select(WageData10, - OCC\_TITLE)  
WageData10 <- mutate(WageData10, YEAR = 2010)  
  
  
WageData09 <- read\_excel("state\_dl.xls" , na = "")  
WageData09 <- select(WageData09, STATE, OCC\_TITLE, TOT\_EMP, H\_MEAN, A\_MEAN, H\_MEDIAN, A\_MEDIAN)   
WageData09 <- filter(WageData09, OCC\_TITLE == "All Occupations")  
WageData09 <- select(WageData09, - OCC\_TITLE)  
WageData09 <- mutate(WageData09, YEAR = 2009)  
  
WageData19 <- WageData19 %>% rename(TOT\_EMP = tot\_emp, H\_MEAN = h\_mean, A\_MEAN = a\_mean, H\_MEDIAN = h\_median, A\_MEDIAN = a\_median)  
  
All\_Data <- rbind(WageData19, WageData18, WageData17, WageData16, WageData15, WageData14, WageData13, WageData12, WageData11, WageData10, WageData09)

After collecting and cleaning the BLS data next I downloaded and cleaned the Job Openings and Labor Turnover Survey(JOLTS). I also began merging data.

All\_Data <- rbind(WageData19, WageData18, WageData17, WageData16, WageData15, WageData14, WageData13, WageData12, WageData11, WageData10, WageData09)  
  
JOLTSData <- read\_excel("jlt\_statedata\_q4\_2019.xlsx" , na = "")

## New names:  
## \* `` -> ...2  
## \* `` -> ...3  
## \* `` -> ...4  
## \* `` -> ...5  
## \* `` -> ...6  
## \* … and 7 more problems

JOLTSData <- JOLTSData[-c(1),]  
JOLTSData <- JOLTSData[-c(1),]  
JOLTSData <- JOLTSData[-c(1),]  
JOLTSData <- JOLTSData[-c(1),]  
JOLTSData <- rename(JOLTSData, "YEAR" = "TOTAL NONFARM, February 2001-December 2019", "STATE" = "...2", "REGION" = "...3", "Job Openings" = "...4", "Hires" = "...5", "Quits" = "...6", "Layoffs & Discharges" = "...7", "Total Separations" = "...8", "Job Openings Rate" = "...9", "Hires Rate" = "...10", "Quits Rate" = "...11", "Layoffs & Discharges Rate" = "...12", "Total Separations Rate" = "...13")  
  
JOLTSData <- JOLTSData %>% mutate(YEAR = replace(YEAR, YEAR == 201905, 2019))  
JOLTSData <- JOLTSData %>% mutate(YEAR = replace(YEAR, YEAR == 201805, 2018))  
JOLTSData <- JOLTSData %>% mutate(YEAR = replace(YEAR, YEAR == 201705, 2017))  
JOLTSData <- JOLTSData %>% mutate(YEAR = replace(YEAR, YEAR == 201605, 2016))  
JOLTSData <- JOLTSData %>% mutate(YEAR = replace(YEAR, YEAR == 201505, 2015))  
JOLTSData <- JOLTSData %>% mutate(YEAR = replace(YEAR, YEAR == 201405, 2014))  
JOLTSData <- JOLTSData %>% mutate(YEAR = replace(YEAR, YEAR == 201305, 2013))  
JOLTSData <- JOLTSData %>% mutate(YEAR = replace(YEAR, YEAR == 201205, 2012))  
JOLTSData <- JOLTSData %>% mutate(YEAR = replace(YEAR, YEAR == 201105, 2011))  
JOLTSData <- JOLTSData %>% mutate(YEAR = replace(YEAR, YEAR == 201005, 2010))  
JOLTSData <- JOLTSData %>% mutate(YEAR = replace(YEAR, YEAR == 200905, 2009))  
JOLTSData <- JOLTSData %>% mutate(YEAR = replace(YEAR, YEAR == 200805, 2008))  
JOLTSData <- JOLTSData %>% mutate(YEAR = replace(YEAR, YEAR == 200705, 2007))  
JOLTSData <- JOLTSData %>% mutate(YEAR = replace(YEAR, YEAR == 200605, 2006))  
JOLTSData <- JOLTSData %>% mutate(YEAR = replace(YEAR, YEAR == 200505, 2005))  
JOLTSData <- JOLTSData %>% mutate(YEAR = replace(YEAR, YEAR == 200405, 2004))  
JOLTSData <- JOLTSData %>% mutate(YEAR = replace(YEAR, YEAR == 200305, 2003))  
JOLTSData <- JOLTSData %>% mutate(YEAR = replace(YEAR, YEAR == 200205, 2002))  
JOLTSData <- JOLTSData %>% mutate(YEAR = replace(YEAR, YEAR == 200105, 2001))  
  
All\_Data2 <- merge(All\_Data, JOLTSData, by = c("STATE", "YEAR"))  
EduAttainment <- read\_excel("Edu\_Attainment.xlsx")  
EduAttainment <- rename(EduAttainment, "YEAR" = "Year")  
EduAttainment <- rename(EduAttainment, "STATE" = "State")  
  
All\_Data3 <- merge(All\_Data2, EduAttainment, by = c("STATE", "YEAR"))

Next I uploaded a document that identified RTW states and created a bianary variable that I used for the regressions 1 = RTW and 0 = Non-RTW

RTWStates <- read\_excel("RTW\_States.xlsx")  
RTWStates <- rename(RTWStates, "STATE" = "State")  
All\_Data4 <- merge(All\_Data3, RTWStates, by = "STATE")  
All\_Data5 <- mutate(All\_Data4, ifelse(YEAR >= Year\_RW, 1, 0))  
All\_Data5 <- rename(All\_Data5, "RTW" = "ifelse(YEAR >= Year\_RW, 1, 0)")  
  
export(All\_Data5, "All\_DataFive.xlsx")  
  
All\_Data6 <- read\_excel("All\_Data5.xlsx")

This next section I used to control for union data.

STATE <- select(RTWStates, STATE)  
UnionMemsByState <- read\_excel("SeriesReport-20200502131616\_b4906a.xlsx")

## New names:  
## \* `` -> ...2  
## \* `` -> ...3  
## \* `` -> ...4  
## \* `` -> ...5  
## \* `` -> ...6  
## \* … and 6 more problems

UnionMemsByState <- UnionMemsByState[-c(1),]  
UnionMemsByState <- UnionMemsByState[-c(1),]  
UnionMemsByState <- UnionMemsByState[-c(1),]  
UnionMemsByState <- rename(UnionMemsByState, "STATES" = "Union affiliation data from the Current Population Survey")  
UnionMemsByState <- rename(UnionMemsByState, "2009" = "...2", "2010" = "...3", "2011" = "...4", "2012" = "...5", "2013" = "...6", "2014" = "...7", "2015" = "...8", "2016" = "...9", "2017" = "...10", "2018" = "...11", "2019" = "...12")  
UnionMemsByState <- cbind(UnionMemsByState, STATE)  
UnionMemsByState <- select(UnionMemsByState, - STATES)  
UnionMemsByState\_L = gather(UnionMemsByState, YEAR, NumberOfUnionMems, - STATE)

This portion contains the last two data sets I used for my regression. After combining everything, R was reading the numbers a texts, because thats how they were saved in Excel when I uploaded to R it gave me error messages. After investigating I found there was no way to fix this with R so I had to export it back to excel and revert them to a number format.

All\_Data7 <- merge(All\_Data6, UnionMemsByState\_L, by = c("YEAR", "STATE"))  
export(All\_Data7, "All\_DataSeven.xlsx")  
All\_Data8 <- read\_excel("All\_Data7.xlsx")

Lastly, for my data upload and cleaning, I uploaded more union data that I also used for the interactive dashboard portion of the paper.

StateUnionMems <- read\_excel("StateUnionMemss.xlsx")  
StateUnionMems <- mutate(StateUnionMems, NumberOfUnionMems = NumberOfUnionMems\*100)  
StateUnionMems <- mutate(StateUnionMems, UnionPercentOfWorkForce = (NumberOfUnionMems/TOT\_EMP)\*100)  
Wage <- select(All\_Data8, YEAR, STATE, H\_MEAN)  
Wage <- filter(Wage, YEAR =="2018")  
StateUnionMems <- merge(StateUnionMems, Wage, by = c("STATE", "YEAR"))

## Results

 After running many regressions, I found the results to be highly mixed. The first one I ran is where I regressed mean hourly wages against all of my independent variables.As one may see all my data fits the model well with a R^2 and adjusted R^2 value of 0.78. In other words, 78% of the variance in my dependent variable can be explained by the variance in my independent variables. All variables are indeed statistically significant at he 0.05 level with the exception of, Hires, Quits, NoHigh, Some and NumberOfUnionMems. From my sample of data, I fail to reject the null hypothesis. I can confidently say that a state having RTW laws can expect to see a 0.64 unit decrease in mean hourly wages, and is statistically significant at the .01 level.

All\_Variables\_Regression <- lm(H\_MEAN ~ RTW + TOT\_EMP + JobOpenings + Hires + Quits + LayoffsAndDischarges + Children + NoHigh + Some + Bachelors + NumberOfUnionMems, data = All\_Data8)  
summ(All\_Variables\_Regression)

## MODEL INFO:  
## Observations: 510  
## Dependent Variable: H\_MEAN  
## Type: OLS linear regression   
##   
## MODEL FIT:  
## F(11,498) = 160.52, p = 0.00  
## R² = 0.78  
## Adj. R² = 0.78   
##   
## Standard errors: OLS  
## ---------------------------------------------------------  
## Est. S.E. t val. p  
## -------------------------- ------- ------ -------- ------  
## (Intercept) 6.71 2.01 3.33 0.00  
## RTW -0.63 0.19 -3.24 0.00  
## TOT\_EMP 0.00 0.00 2.83 0.00  
## JobOpenings 0.00 0.00 0.34 0.73  
## Hires 0.00 0.01 0.22 0.83  
## Quits -0.01 0.01 -1.12 0.27  
## LayoffsAndDischarges -0.02 0.01 -2.17 0.03  
## Children 18.83 5.05 3.73 0.00  
## NoHigh 1.21 5.30 0.23 0.82  
## Some -4.85 4.16 -1.17 0.24  
## Bachelors 56.29 2.44 23.10 0.00  
## NumberOfUnionMems -0.00 0.00 -0.45 0.65  
## ---------------------------------------------------------

 For robustness testing, that is ensuring that my results are consistent across different but related dependent variables in this case income metrics, I regressed the same variables but this time on mean annual salaries.As with my previous model the data fits well with both R^2 and adjusted R^2 values equal to 0.78, with 78% of the variance in my independent variables explaining the variance in my dependent variables. All variables are indeed statistically significant at he 0.05 level with the exception of JobOpenings, Hires, Quits, NoHigh, Some and NumberOfUnionMems. According to these results, states that have passed RTW laws can expect to see a 1,313.98 unit decrease in the mean annual salary. The results are statistically significant. From this regression results I fail to reject the null hypothesis. There is not enough evidence to conclude that RTW states have higher wages than Non-RTW states.

All\_Variables\_Regression2 <- lm(A\_MEAN ~ RTW + TOT\_EMP + JobOpenings + Hires + Quits + LayoffsAndDischarges + Children + NoHigh + Some + Bachelors + NumberOfUnionMems, data = All\_Data8)  
summ(All\_Variables\_Regression2)

## MODEL INFO:  
## Observations: 510  
## Dependent Variable: A\_MEAN  
## Type: OLS linear regression   
##   
## MODEL FIT:  
## F(11,498) = 160.51, p = 0.00  
## R² = 0.78  
## Adj. R² = 0.78   
##   
## Standard errors: OLS  
## -----------------------------------------------------------------  
## Est. S.E. t val. p  
## -------------------------- ----------- ---------- -------- ------  
## (Intercept) 13969.56 4190.48 3.33 0.00  
## RTW -1313.98 405.32 -3.24 0.00  
## TOT\_EMP 0.00 0.00 2.84 0.00  
## JobOpenings 3.38 9.99 0.34 0.73  
## Hires 3.27 14.88 0.22 0.83  
## Quits -22.80 20.45 -1.11 0.27  
## LayoffsAndDischarges -38.65 17.80 -2.17 0.03  
## Children 39162.54 10502.87 3.73 0.00  
## NoHigh 2523.18 11025.79 0.23 0.82  
## Some -10080.87 8649.93 -1.17 0.24  
## Bachelors 117087.86 5068.84 23.10 0.00  
## NumberOfUnionMems -0.17 0.37 -0.45 0.65  
## -----------------------------------------------------------------

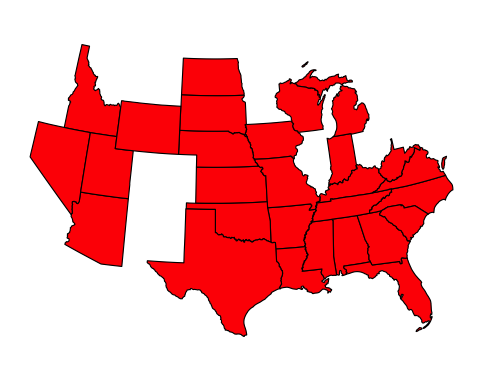
This next section uploads some data to aid in visualization

usmap::plot\_usmap("states", fill = "red", alpha = 1, include = c("AL", "AZ", "AR", "FL", "GA", "ID", "IN", "IA",  
 "KS", "KY", "LA", "MI", "MS", "MO", "NE", "NV",   
 "NC", "ND", "OK", "SC", "SD", "TN", "TX", "UT",  
 "VA", "WV", "WI", "WY"))

## Warning: Use of `map\_df$x` is discouraged. Use `x` instead.

## Warning: Use of `map\_df$y` is discouraged. Use `y` instead.

## Warning: Use of `map\_df$group` is discouraged. Use `group` instead.

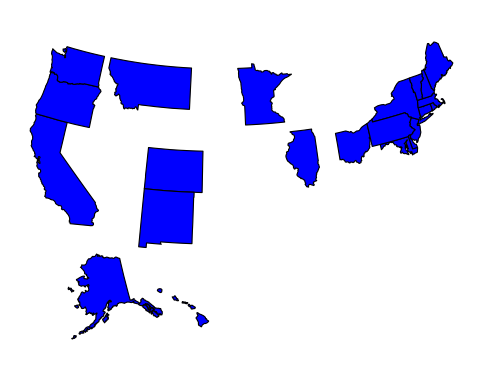


usmap::plot\_usmap("states", fill = "blue", alpha = 1, include = c("AK", "CA", "CO", "CT", "DE", "DC", "HI", "IL",  
 "ME", "MD", "MA", "MN", "MT", "NH", "NJ", "NM",   
 "NY", "OH", "OR", "PA", "RI", "VT", "WA"))

## Warning: Use of `map\_df$x` is discouraged. Use `x` instead.

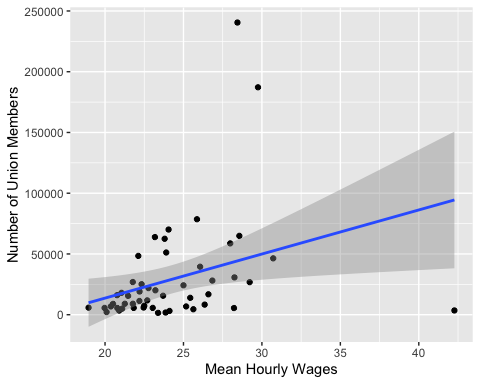
## Warning: Use of `map\_df$y` is discouraged. Use `y` instead.

## Warning: Use of `map\_df$group` is discouraged. Use `group` instead.



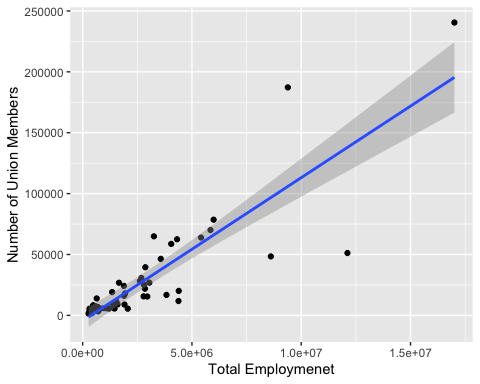
ggplot(StateUnionMems, aes(x=H\_MEAN, y=NumberOfUnionMems)) + geom\_point() +  
 geom\_smooth(method="lm") + labs(x="Mean Hourly Wages", y="Number of Union Members")

## `geom\_smooth()` using formula 'y ~ x'



ggplot(StateUnionMems, aes(x=TOT\_EMP, y=NumberOfUnionMems)) + geom\_point() +  
 geom\_smooth(method="lm") + labs(x="Total Employmenet", y="Number of Union Members")

## `geom\_smooth()` using formula 'y ~ x'



## Implications

 For other researches investigating this matter, I would suggest to control for other cost of living expenses when investigating wages. It is believed that states with higher wages have higher living expenses. Controlling for this may change the results drastically. Another variable I would suggest to control for are percentage shares of industry within a state, i.e. percentage of GDP from manufacturing, services, private, public, etc. This may also change regression results drastically.

## Conclusion

 This report has been very telling, as I said, from the data I collected I have failed to reject the null hypothesis of H\_o: RTWStateWagesNonRTWStateWages. I hope that this research contributes to the existing literature in an effort to improve the incomes and lives of all in the United States.