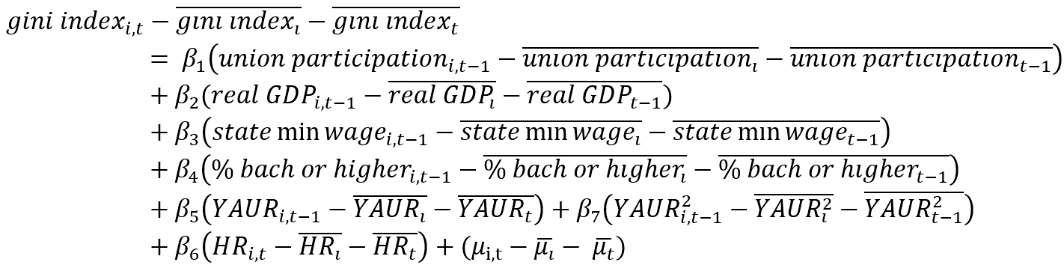
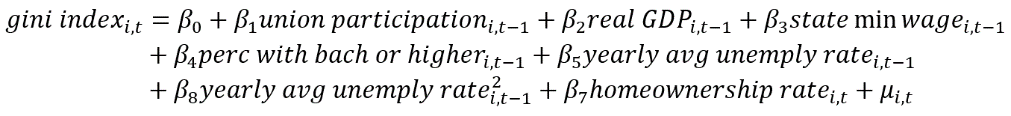
Methodology

Our empirical approach is summarized by follow two equations.

1.

2.

In both equation one and two we regress Gini index on the percent of the work force represented by a union for state I at time t while controlling for several lagged and non-lagged variables we believe also influence a state’s income distribution. The difference between the two equations is that equations 1 contains state and time fixed effects, and equation two is a pooled OLS that does not account for state of time heterogeneity. Since we have a panel data, there is a good chance that we have heterogeneity bias between states and years. If so, pooled OLS would not be a good estimation technique because the covariance between our independent variables and the unobserved heterogeneity will not be zero. That will make POLS estimates biased and inconsistent.

To decide if we should include fixed effects in our model performed an F-test for individual and time effects. The null hypothesizes is that we have homogeneous data and have no need for fixed effects. The F-test is statistically significant result with a p-value of less than 2.22e-16. We reject the null hypothesis and conclude that we have heterogeneity in our data, and fixed effects is the correct specification.

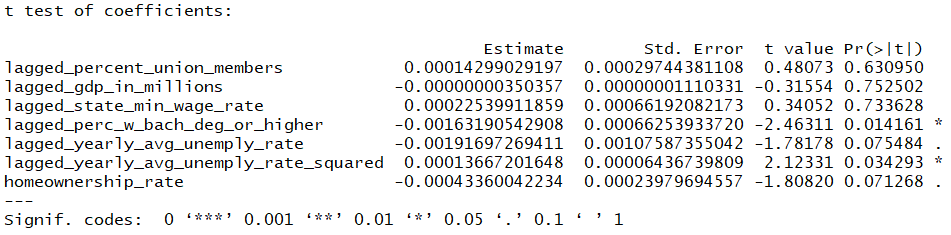
Next, we tested for heteroskedastic and serial correlation. We preformed the Breusch-Pagan test for heteroskedastic and the Breusch-Godfrey test for serial correlation of the error term. The null hypothesis in the tests is that heteroskedastic and serial correlation are not present. We found statistically significant results in both test with p-values less than 0.05. We reject the null hypothesis and conclude that heteroskedastic and serial correlation are present. We obtained Arellano robust standard errors for correct for these issues.

Finally, we checked the presences of multicollinearity in our model. We calculated the variance inflation factor for the fixed effects model. We did not find evidence of multicollinearity, as none of our variables at an VIF above 2, expect for lagged average unemployment rate and lagged average unemployment rate squared. We tested the VIF of our model both including and excluding lagged average unemployment rate squared and found that it is only the VIF of lagged average unemployment rate was below 2 when lagged average unemployment rate squared was not present.

We are unable to address whether the exogeneity assumptions holds for our model. We were unable to find a suitable interment variable for union participation rate to test if it is exogenous.

Results

Our results with the robust standard errors are below.



We did not find evidence that state and time fixed lagged union participation influences on the distribution of income. However, we did find evidence that lagged percent with a bachelor’s degrees, lagged average yearly unemployment, and homeownership rate influence the distribution of income.