Impact of Teachers' Union Concentration on Student Academic Performance

Summer Will

Pomona College

ECON057: Economic Statistics

Dr. Michelle Zemel

January 3, 2021

Introduction

Education budgets go towards many aspects of learning, including teachers' salaries and student resources. Teachers' unions, like other types of unions, were begun in order to give their members power to fight for appropriate wages, and in order to let workers maintain the ability to exercise their rights. A portion of money allocated to the largest teachers' unions of today, such as the National Education Association (NEA) and the American Federation of Teachers (AFT), goes toward federal lobbying practices (Cowen & Strunk, 2015). These unions use Collective Bargaining Agreements (CBAs) to organize and govern the rights of its members. While teachers' unions were created out of an effort to raise up the rights of teachers and by extension, students, it can be argued that education funds can be better allocated away from administration and lobbying, and towards students and teachers. In many cases, education budgets (local or state) need more money for the sake of furthering the academic achievement of students. In order to make sure education fund increases reach teachers and students, however, CBAs much be investigated and refined so that the allocation is fair. In order to determine the efficacy of teachers' unions in the modern age, the question must be asked: how effective are teachers' unions? Namely, are they effective enough to defend their administrative and political expenses? This paper will attempt to find any causality between the power teachers' unions in a given area and measures of academic success within that area. It is hypothesized that teachers' unions will have a negative impact on students' academic success.

Data

Union data and education data do not coincide well within detailed, publicly available datasets such as the National Center for Education Statistics or the American Community

Survey. Educational statistics exist primarily by state and district, while the most readily-available measures of union concentration in a given area are available by county. In order to compare academic and union statistics by county, union data was cross-referenced from the Current Population Survey (CPS) with the only state found to have county-wide academic achievement data: California. The California Assessment of Student Performance and Progress (CAASPP) administers Smarter Balanced Assessments yearly, and assesses students by a consistent, standardized measure of performance.

The CPS surveys thousands of people across the United States every year, collecting data on the surveyed individual's education, household characteristics, occupational characteristics, and more (Flood, Sarah et al., 2020). The Integrated Public Use Microdata Series compiles CPS data, offering samples by year or month, with each sample offering comprehensive and varying data--this paper uses survey data from 2014 to 2018, as the 2019 and 2020 samples lacked necessary variables about an individual's educational attainment.

This paper discusses "union concentration", which hereby refers to how much union involvement there is in a given county, among elementary, primary, and secondary teachers. The CPS offers one measure of union involvement that is on a scale from 0 to 3, with 0) NIU, 1) no union coverage, 2) member of labor union, and 3) covered by union but not a member. This paper's data analysis on union concentration takes the mean survey response by county on a scale from 1 to 2, no union coverage to member of labor union. This was done by the collapse tool in STATA: 2014 to 2018 samples were scaled down to include only elementary, primary, and secondary teachers, and those teachers who either have no union coverage or are a member of a labor union. The remaining observations were then collapsed into union mean by county, resulting in a collection of counties with their respective union concentrations on a scale from 1 to 2.

CAASPP Smarter Balanced Assessments gauge academic performance on geographic bases by state, district, and county, and topical bases by English Language Arts, science, and mathematics. In this paper, academic performances of select California counties will be analyzed by the percentage of students who “met or exceeded” English Language Arts standards and students who “met or exceeded” mathematics standards. This analysis will be among all grade levels for which the data is offered, 3rd through 11th grade. The data will include all of public, private, and charter schools (California, 2020).

As discussed, educational data and union data are difficult to compare due to the geographies by which each is most often based. This paper seeks to utilize a measure of general academic circumstance beyond academic performance, since this data was only available for select California counties. The CPS does not survey drop-out rates, nor does it ask a question related, specifically, to dropping out. In order to create a variable that measures drop-out rates, three different educational variables were used from a 2018 CPS sample: educational attainment (educ99), currently attending regular school (edatt), and whether or not the individual was enrolled in school last October (edattly). The sample data was scaled down to: those who had attained some measure of education up through 12th grade, but had not graduated high school or received a GED; those who are were not then-currently attending school; and those who were enrolled in school last October. Theoretically, this would create a 2018 pool of individuals who dropped out, since they had attended school the previous October, were not then-currently in school, and who had not yet graduated. See Discussion for ways in which this particular measure of student drop-outs may not have been ideal. Table 2 measures union concentration of the available counties from 2014-2018 data. This table will be used to compare union concentration and drop-out rates from 2018. Not all CPS counties had student “drop-outs”, so the available CPS counties in the sample data have been reduced to the ones that displayed any amount of drop-outs.

In order to properly regress academic achievement or drop-out rates on union concentration, confounding variables that lead to the outcome variables must be accounted for. This paper has accounted for household income, property taxes, and race. Since union concentration was calculated using teacher responses the control variables were calculated separately on a broader basis—the union counties and control counties were then matched.

Table 1 displays the cross-referenced union concentration and academic performances of select California counties. Union, household income, property tax, and race data come from 2018 samples, since the academic performance was measured from the 2018-2019 school year. There were 24 counties total which had both academic achievement and union data.

county (FIPS)	union	ELA (%)	math (%)	hhincome (U	proptax (USD
6061	1	65.47	54.04	110502.665	2122.51648
6089	1.4	48.81	39.44	80113.5979	1823.16495
6111	1.42857	48.4	36.88	104454.672	2463.27356
6059	1.48387	59.69	50.44	121215.99	1924.45527
6075	1.5	56.17	49.48	167425.062	2459.87115
6083	1.5	47.22	36.21	100389.838	1817.05051
6097	1.5	50.4	37.94	110678.222	2790.75309
6073	1.54545	56.76	45.05	110832.43	1984.19961
6095	1.57143	45.57	33.88	91032.2344	1327.03646
6081	1.6	61.54	53.37	227572.902	3089.53517
6001	1.64286	56.84	48.98	144444.101	3479.21304
6037	1.64762	50.43	39.11	102856.291	1911.11268
6077	1.71429	43.45	30.92	84972.4297	2072.40849
6067	1.76923	48.85	37.88	96064.5668	1567.51542
6029	1.82353	43.44	28.83	72404.0791	1034.51276
6107	1.875	43.09	29.28	79184.5714	1360.01071
6019	1.91667	48.04	36.51	81587.7323	1107.6936
6007	2	49.1	35.99	73164.9279	1820
6017	2	60.96	50.19	113977.515	1972.74242
6053	2	39.77	27.4	88065.1498	1766.48792
6079	2	56.9	45.53	117339.924	2504.71739
6087	2	47.16	35.82	103981.445	2525.64384
6099	2	42.81	29.93	83937.8639	3567.13613
6113	2	51.33	39.62	84457.3711	1023.90722

Table 1

county	union	#drop-out	hhincome	proptax
37179	1	4	106953.133	1163.14221
48061	1.05769	2	56222.7913	1568.88679
37001	1.05882	8	76792.2913	1052.4849
12111	1.07143	2	70094.6327	1949.5496
47165	1.07143	2	102325.531	944.633663
48215	1.09333	4	51572.3665	1560.14192
4019	1.1	4	68289.0939	951.453049
4027	1.11111	2	67281.317	1012.25313
37133	1.125	2	66733.4892	877.33665
37191	1.125	4	60994.9039	672.81146
51153	1.22727	2	124007.372	3120.29044
12103	1.26582	4	82769.6932	1673.74033
4013	1.27907	10	86822.7141	964.815965
12086	1.33557	6	76319.6692	1989.30922
22103	1.36364	2	116137.043	1069.86017
12099	1.36634	4	94411.6619	2106.46455
6083	1.38095	4	102216.775	1702.11933
11001	1.38692	8	125123.37	1688.36337
12069	1.3871	6	80451.8573	2009.47881
35045	1.4	2	67193.5534	818.510907
1003	1.44737	6	69511.446	598.526071
35001	1.44776	2	86922.3467	1249.26441
8059	1.45455	2	114501.901	1404.08345
13139	1.45455	2	84397.4674	842.98427
18039	1.45455	4	98082.2738	1344.59354
31055	1.45455	4	99968.0935	2201.15775
19163	1.52381	8	106788.992	2238.87074
24031	1.53448	2	151877.647	3045.13433
10003	1.54726	8	99325.3054	3534.27706
54039	1.5614	4	79272.6403	752.349564
32003	1.58683	4	78537.8059	1259.0669
24025	1.6087	8	119450.815	4216.89168
15003	1.61682	2	114729.302	1071.27413
6037	1.62025	10	91916.2877	1913.9271
36047	1.63309	4	85459.9129	1275.25249
6081	1.63636	2	175491.111	3480.00175
26163	1.64384	2	71761.6214	2028.09828
41029	1.64706	4	84665.3008	1269.19036
6073	1.67836	4	100942.448	2061.74917
6053	1.69565	2	84369.9494	2015.55015
6111	1.71739	6	108519.293	2358.16747
36119	1.71739	2	144990.599	5855.36241
41039	1.72414	2	75336.397	1705.44265
25021	1.72917	2	153498.195	4413.4699
6019	1.73846	2	75867.3027	1353.81298
34005	1.75	2	119465.523	4233.24927
33011	1.76522	2	117777.008	4106.9745
34039	1.77419	2	120762.468	5546.79307
36059	1.78235	4	152155.533	7653.57918
53057	1.78571	2	94881.1658	2401.92513
36085	1.84375	2	101332.762	3599.61105
36087	1.88889	2	130244.722	6574.25343
42071	1.88889	2	96664.7321	3262.33656
36103	1.91791	2	118805.745	6868.51295
36071	1.92105	2	115507.207	5649.23596
6017	2	2	89525.0119	2267.94328

Table 2

Table 2 displays the mean number of drop-outs, household income, and property tax per county. There are 56 observations.

```
. regress ELA union hhincomeUSD proptaxUSD
```

Source	SS	df	MS
Model	539.883134	3	179.961045
Residual	541.770266	20	27.0885133
Total	1081.6534	23	47.0284087

Number of obs	=	24
F(3, 20)	=	6.64
Prob > F	=	0.0027
R-squared	=	0.4991
Adj R-squared	=	0.4240
Root MSE	=	5.2047

ELA	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
union	-6.844042	4.308562	-1.59	0.128	-15.83154	2.143461
hhincomeUSD	.0001265	.0000401	3.16	0.005	.0000429	.0002101
proptaxUSD	-.0009361	.0018855	-0.50	0.625	-.0048692	.002997
_cons	51.07881	9.184331	5.56	0.000	31.92063	70.23699

Regression 1a

```
. regress math union hhincomeUSD proptaxUSD
```

Source	SS	df	MS
Model	902.309049	3	300.769683
Residual	610.357285	20	30.5178642
Total	1512.66633	23	65.7681014

Number of obs	=	24
F(3, 20)	=	9.86
Prob > F	=	0.0003
R-squared	=	0.5965
Adj R-squared	=	0.5360
Root MSE	=	5.5243

math	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
union	-8.007891	4.573165	-1.75	0.095	-17.54734	1.531563
hhincomeUSD	.0001655	.0000425	3.89	0.001	.0000767	.0002542
proptaxUSD	-.0009674	.0020013	-0.48	0.634	-.005142	.0032072
_cons	37.75928	9.74837	3.87	0.001	17.42453	58.09402

Regression 1b

```
. regress dropout union hhincome proptax

      Source |       SS           df       MS      Number of obs   =        56
-------------+----------------------------------   F(3, 52)        =      0.87
       Model |  13.4584806         3  4.48616021   Prob > F        =    0.4614
    Residual |  267.398662        52  5.14228197   R-squared       =    0.0479
-------------+----------------------------------   Adj R-squared   =   -0.0070
       Total |  280.857143        55  5.10649351   Root MSE        =    2.2677

------------------------------------------------------------------------------
     dropout |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       union |  -1.324015   1.457555    -0.91   0.368    -4.248813    1.600784
    hhincome |   1.25e-06   .0000156     0.08   0.937    -.0000301    .0000326
     proptax |  -.0001254   .0002737    -0.46   0.649    -.0006746    .0004237
       _cons |   5.816735   2.167761     2.68   0.010     1.466805    10.16667
------------------------------------------------------------------------------
```

Regression 2

We observe negative relationships in the primary sample coefficient for each of the three regressions. For Regression 1, the correlation points to a negative effect of teachers' union concentration on academic performance, but for Regression 2, the correlation points to a positive effect of union concentration on drop-outs (increase in union means less drop-outs). However, all regressions present a p-value that prevents rejecting the null hypothesis. When analyzing the estimated coefficients, other x-variables are held constant, allowing more focused control over one variable's effect on another.

Methodology

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \varepsilon_i$$

Above is the statistical model for (multivariate) Regression 1, where:

y_i = union concentration of county *i*

x_{1i} = mean academic performance of county *i*

**where academic performance is calculated by ELA in Regression 1a and math in Regression 1b*

x_{2i} = mean household income of county *i*

x_{3i} = mean property tax of county *i*

β_0 = y-intercept

β_1 = the relationship (slope) between union concentration and academic performance

β_2 = the relationship (slope) between union concentration and mean household income

β_3 = the relationship (slope) between union concentration and mean property tax

ε_i = error of observation *i*

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \varepsilon_i$$

Above is the statistical model for Regression 2, where:

y_i = union concentration of county *i*

x_{1i} = mean number of student drop-outs in county *i*

x_{2i} = mean household income of county *i*

x_{3i} = mean property tax of county *i*

β_0 = y-intercept

β_1 = the relationship (slope) between union concentration and number of student drop-outs

β_2 = the relationship (slope) between union concentration and mean household income

β_3 = the relationship (slope) between union concentration and mean property tax

ε_i = error of observation *i*

The regression tool on STATA shows results of hypothesis testing. Shown on Regressions 1 and 2 are the number of standard deviations away from the null mean in the t-distribution that there would be given the null that the results were by chance. In Regression 1a, the results were 1.59 standard deviations away from the t-distribution mean, giving them a p-value of 12.8 percent. We fail to reject the null hypothesis. In Regression 1b the results were 1.75 standard deviations away from the t-distribution mean, giving them a p-value of 9.5 percent. We fail to reject the null hypothesis. In Regression 2 the results were .91 standard deviations away from the t-distribution mean, giving them a p-value of 36.8 percent. We fail to reject the null hypothesis.

The t-distribution was used to calculate the p-values for Regressions 1 and 2, as opposed to the normal distribution. The t-distribution is used when variance is estimated, not known, which is the case in the regressions above. In addition, there were only 24 observations for Regression 1, meaning that particular sample did not fall under the Central Limit Theorem and therefore not the Normal distribution, even if the variance were known. Conceptually, the estimated relationship for Regression 1a between union concentration and ELA, is, for every increase of 1 in union concentration there was a decrease of about 7 percentage points. This negative correlation points to a negative effect of teachers' union concentration on academic performance, but since we failed to reject the null hypothesis, we cannot draw conclusions from this relationship. Similar situations follow with Regression 1b and Regression 2; the former showed another negative correlation which cannot be confirmed due to failure to reject the null. Regression 2 conceptually pointed the opposite direction, showing a decrease in drop-outs for any increase in teachers' union concentration. Again, we failed to reject the null, so no true conclusions can be drawn.

Results

Regression 1a shows the result of estimating the relationship between union concentration and English Language Arts proficiency in a given California county. Let the null hypothesis for this regression analysis be that there is no relationship between these two variables. Even though we observed a negative b_1 value indicating a negative relationship between increased union concentration and academic performance, we fail to reject the null, given a p-value of 12.8 percent.

Regression 1b shows the result of estimating the relationship between union concentration and math proficiency in a given California county. Let the null hypothesis for this regression analysis be that there is no relationship between these two variables. Even though we observed a negative b_1 value indicating a negative relationship between increased union concentration and academic performance, we fail to reject the null, given a p-value of 9.5 percent.

Regression 2 shows the result of estimating the relationship between union concentration and the number of student drop-outs in a given county. Let the null hypothesis for this regression analysis be that there is no relationship between these two variables. Even though we observed a negative b_1 value indicating a negative relationship between increased union concentration and student drop-outs, we fail to reject the null, given a p-value of 36.8 percent.

In regression analysis, the more x-variables that were controlled, the higher the p-value became for each regression. For example, the p-value for a simple regression of union concentration on math proficiency was far less when household income and property taxes were not accounted for. This indicates that household income and property taxes are important confounding variables in the relationship between academic achievement and union concentration. Property taxes have a profound connection to quality of school, since district budgets come directly from property taxes in the majority of districts, so it is not surprising that this variable was confounding. Since property taxes stem mostly from household income, it follows that household income would be a strong confounding variable as well. Having controlled for these two variables, we then failed to reject the null hypothesis. The most indicative result of adding these particular x-variables was the large R-squared value (for Regression 1). We observe an R-squared value of around .5 for Regression 1, which is far larger than is usually to be expected for economic analysis.

Discussion

The aim for this paper was to determine a relationship between some measures of academic success and union concentration. Issues were encountered when searching for data in each of these topics that matched geographically. Educational performance data is available within school districts. Union data is readily available from the Current Population Survey, which has geographical data on many different regions, including county and state, but not district. This discrepancy proved difficult to overcome, as available data across each topic that matched regionally was scant, if at all existent.

Stemming from difficulty finding matching data was the haphazard nature of the drop-out variable. The fact that the drop-out count was an even number—and often the number 2—for each district might show some trend in the data that was unintentional. This measure of a drop-out count was far less than ideal, as it involved a circuitous combination of existing CPS variables instead of an actual survey question about dropping out, so it is somewhat expected that there may have been confounding elements. The results reflect the odd raw drop-out count outcome, since the p-value was 36.8 percent and the R-squared value was .0479, which indicates a sample filled with error. It is possible that the drop-out variable involved consistent measurement error.

The measure of union concentration used in this paper was more extensive than other variables involved, but could have been more comprehensive. One researcher (Moe, 2009) measured union power by creating a multifaceted gauge of Collective Bargaining Agreement strength after reviewing hundreds of CBAs and different constraints teachers faced under them. There are a multitude of elements involved in teachers' union memberships, and the power of teachers' unions becomes impactful in particular scenarios that may be better accounted for with a different measure of union concentration than was used in this paper.

Education is a critical part of a country's success, so it is disappointing to not have been able to come to any concrete conclusions. The lack of causality is, however, understandable given the small amount of fitting data across education and unions. Moving forward this paper would benefit from finding more data on education by county, or to find more data on unions by school district. The relationship between teachers' unions and academic well-being has been found to a reliable extent by other studies, but the true challenge is figuring out what to do with the information. More research ought to be done on exactly what the effect of altering contract restrictiveness would be on academic performance; if that is discovered, then action can be taken in legislation that may improve the education and lives of students across the country.

References

California Department of Education, 2021. California Assessment Of Student Performance And Progress*. Sacramento, CA: California Department of Education.*

Cowen, J. M., & Strunk, K. O. (2015). The impact of teachers' unions on educational outcomes: What we know and what we need to learn. *Economics of Education Review*, *48*, 208–223. https://doi.org/10.1016/j.econedurev.2015.02.006

Flood, Sarah, King, Miriam, Rodgers, Renae, Ruggles, Steven, & Warren, J. Robert. (2020). *Integrated Public Use Microdata Series, Current Population Survey: Version 8.0* (8.0) [Data set]. Minneapolis, MN: IPUMS. https://doi.org/10.18128/D030.V8.0

Moe, T. M. (2009). Collective Bargaining and the Performance of the Public Schools. *Midwest Political Science Association*, *53*(1), 156–174.

Strunk, K. O., & McEachin, A. (2011). Accountability Under Constraint: The Relationship Between Collective Bargaining Agreements and California Schools' and Districts' Performance Under No Child Left Behind. *American Educational Research Journal*, *48*(4), 871–903. https://doi.org/10.3102/0002831211401006