

# State Finance, Partisan Control and Citizen Outcomes: What's the Connection?

Andrew Pederson

2022-29-08

## Introduction

In the U.S.' two-party system, increased polarization has renewed attention to 'Red' vs. 'Blue' states, exemplified by the feud between Florida's Republican Governor DeSantis and California's Democratic Governor Newsom. Each argues that their state represents a clear choice between distinct policies and outcomes, and the Economist recently argued that this trend would increasingly [shape state policies](#) in the future. Is this case? If a different political party controls state governments, do we see different policies and associated outcomes? To what extent do these policy differences significantly affect outcomes for citizens?

In June of 2021, [CNN's Fareed Zakaria reported](#) on a [WalletHub analysis](#) that claims to show just that:



## No Easy Answers

While the WalletHub analysis simplifies some complex data sets into insights effectively, there are some issues. Visually, the graph's y axis is inverted, which gives the misleading impression that more 'Red' states have better services, when the opposite is true. While the axis label clarifies this, it's not natural to look at the lower left quadrant to find the best performers.

The methodology also raises questions. States are coded according to how they voted in the 2020 Presidential election rather than which party controls state government, and states look very different depending on whether we look at the legislature, governorship or both. That the 2020 election was very close and produced some upsets doesn't help; the Republican stronghold of Georgia going to Biden, for example.

It also doesn't make sense to calculate a per capita tax amount by dividing total state income tax revenues by individuals 18 or older nor to exclude local taxes and federal transfers. There are other taxes (sales and use, licensing, etc.) that will be borne by much wider tranches of society, and just because someone is 18 or older is not a guarantee that they are employed nor earn enough to pay taxes. For example, the [Institute on Taxation and Economic Policy](#) regularly reviews how 'equitable' state tax policies are. Though many states may raise significant amounts of tax revenue per capita, those who actually pay the taxes can be concentrated among high-earners, often referred to as 'progressive' tax structures. Looking at this data, average earners in Texas may actually [pay more taxes](#) than in California, even though California technically has much higher income tax rates. Tax revenue sources vary from state to state and may or may not reflect that state's primary revenue source nor what a typical resident can expect to pay to the government as a share of income.

Instead, we can use the [tax burden](#) to estimate the total amount of taxes paid relative to that state's share of the net national product and recreate a version of WalletHub's analysis. In lieu of the 2020 Presidential vote, states are coded according to which political parties controlled its Executive and Legislative Branches in 2020: R for Republican, D for Democratic and S for split control between the two parties.

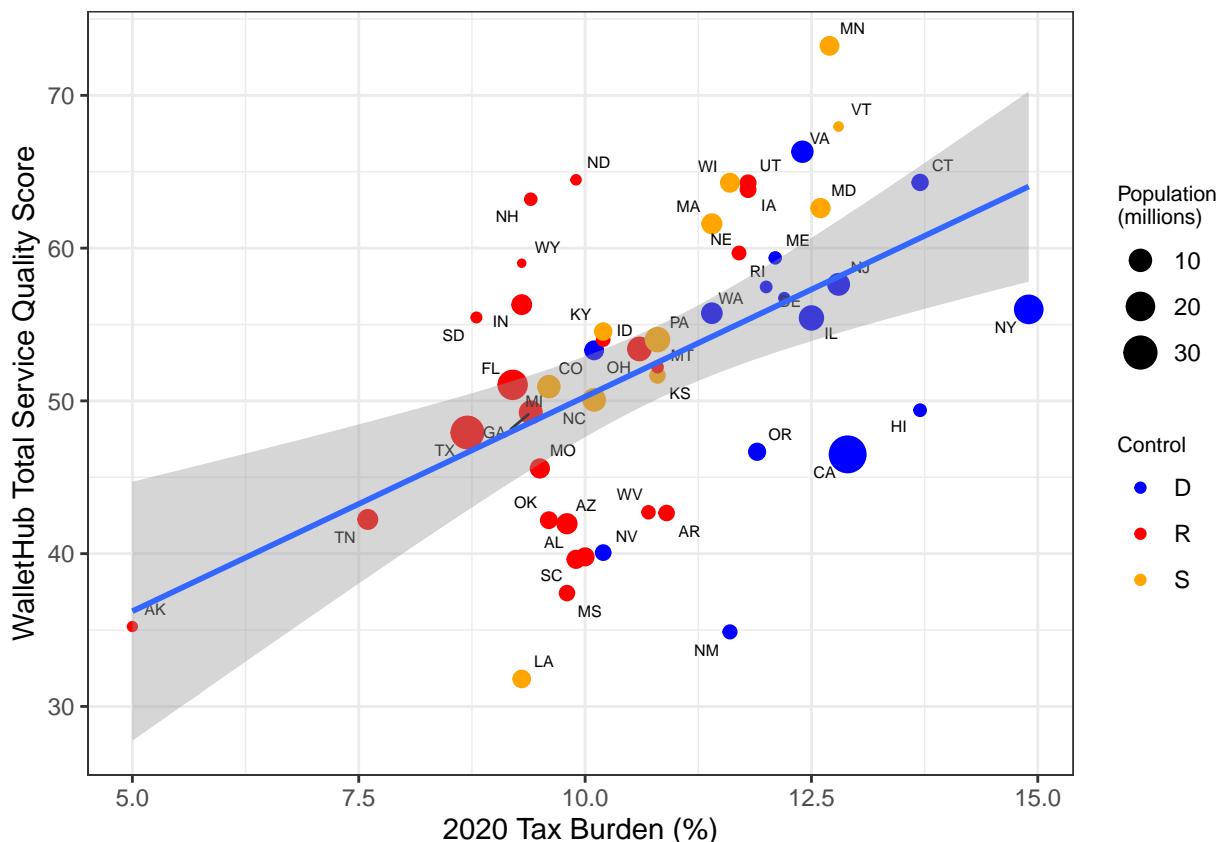


Table 1: Tax Burden’s Effect on Quality Score: Linear Estimate

Predictor	B	SE	t	p
Intercept	22.19	7.622	2.91	0.005
Tax Burden	2.81	0.697	4.03	<0.001

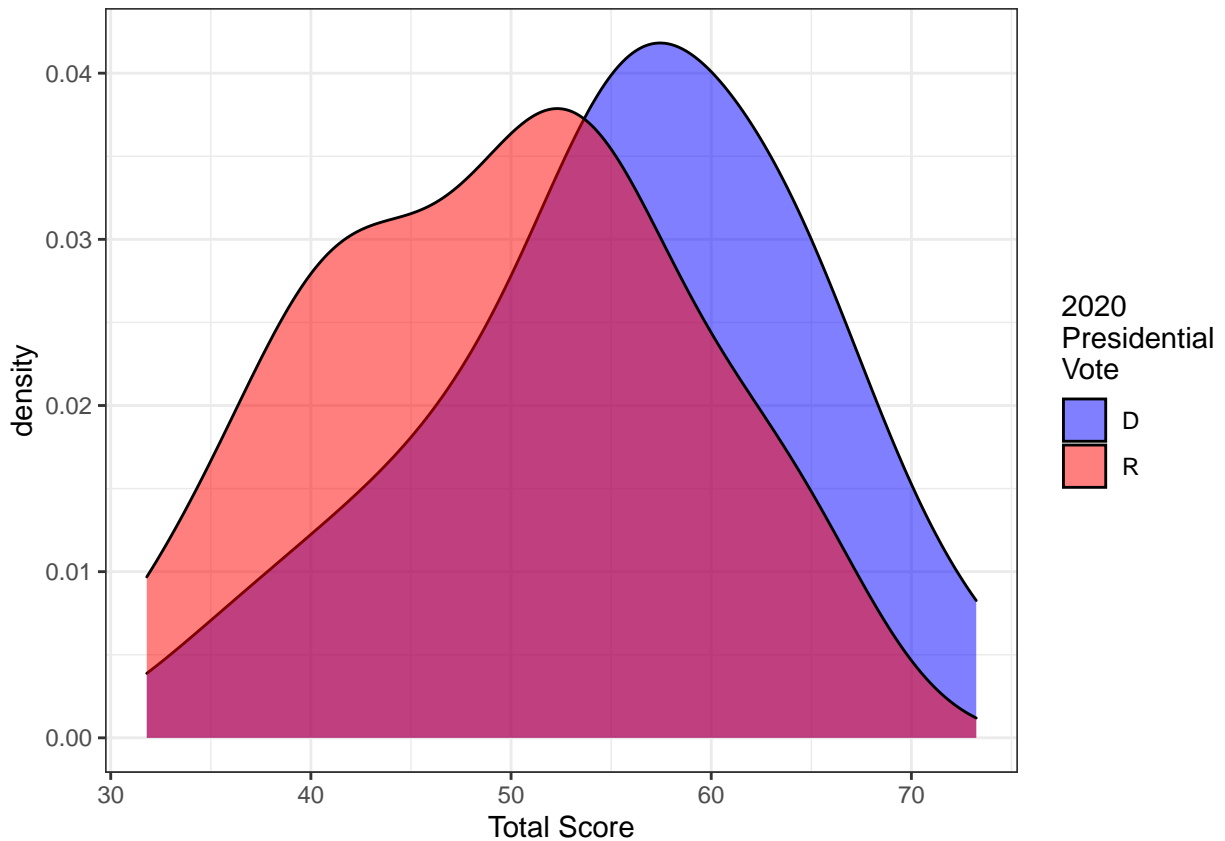
Surprisingly, more so for reasons we’ll see below, a higher tax burden correlates with a better rank in the WalletHub methodology. There is significant variation unrelated to party control, though there are clusters of Republican states with lower tax burdens and Democratic states with higher tax burdens which are statistically significant. It’s also worth noting that throughout this analysis, two of the largest Democratic run states, CA and NY, seem to be outliers, in that their increased spending does not often reflect the better outcomes we might expect.

Table 2: Party Control’s Effect on Tax Burden: Tukey’s HSD

Predictor	Comparison	Null	Estimate	Low	High	p
Party Control	R-D	0	-2.56	-3.624	-1.49	<0.001
Party Control	S-D	0	-1.21	-2.500	0.08	0.069
Party Control	S-R	0	1.34	0.163	2.53	0.022

Another note is that a high ROI score doesn’t mean that the quality of services is very good, despite being efficient. For example, although Missouri ranks 38th in service quality, it ranks 5th for ROI just by dint of not spending very much. Similarly, FL ranks 30th in service quality and 2nd for ROI. CA is 37th in service quality, and though its raw service quality score of 46.5 is not far from Florida at 51.05 or Texas at 47.93, its ROI rank is 49 because it collects significantly more income tax revenue, albeit from a smaller population of higher earning taxpayers.

Statistically, the only significant result we find from the original WalletHub data is that states that voted Democratic in the 2020 Presidential election generally have better services.



Welch Two Sample t-test

```
data: ranks$Total.Score by ranks$X2020.Vote
t = 2.435, df = 47.51, p-value = 0.0187
alternative hypothesis: true difference in means between group D and group R is not equal to 0
95 percent confidence interval:
 1.103244 11.573680
sample estimates:
mean in group D mean in group R
    55.81500     49.47654
```

However, as the next section will show, actual control of the Executive and Legislative Branches is not correlated with better quality services.

## Who's in Office?

Partisan results from a single election is a poor proxy for government control, since some purple states may vote for a Democratic President while the Republican Party still controls the Legislature and/or Governor's office. To examine this, we added three new variables to indicate which party controls each of the Legislative and Executive Branches, and whether each state's Executive and Legislative branches were both controlled by one party or split between the two.

When we look at the Legislature, Governorship and overall control, results are muddy, with the only statistically significant difference between the three groups being that split legislature states provide better services than Republican legislature states, and with only two split legislatures, this conclusion isn't reliable either.

## Legislature

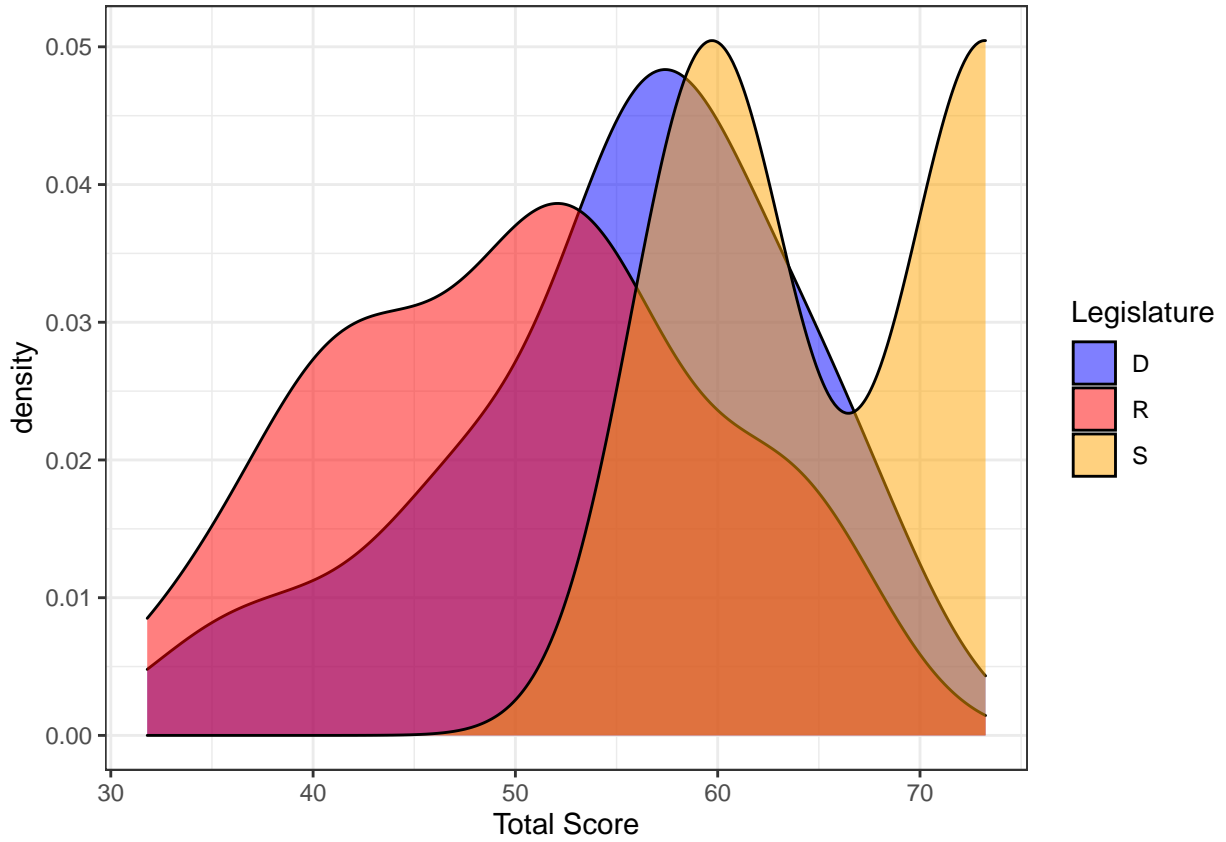


Table 3: Party Legislative Control's Effect on Quality Score: Linear Estimate

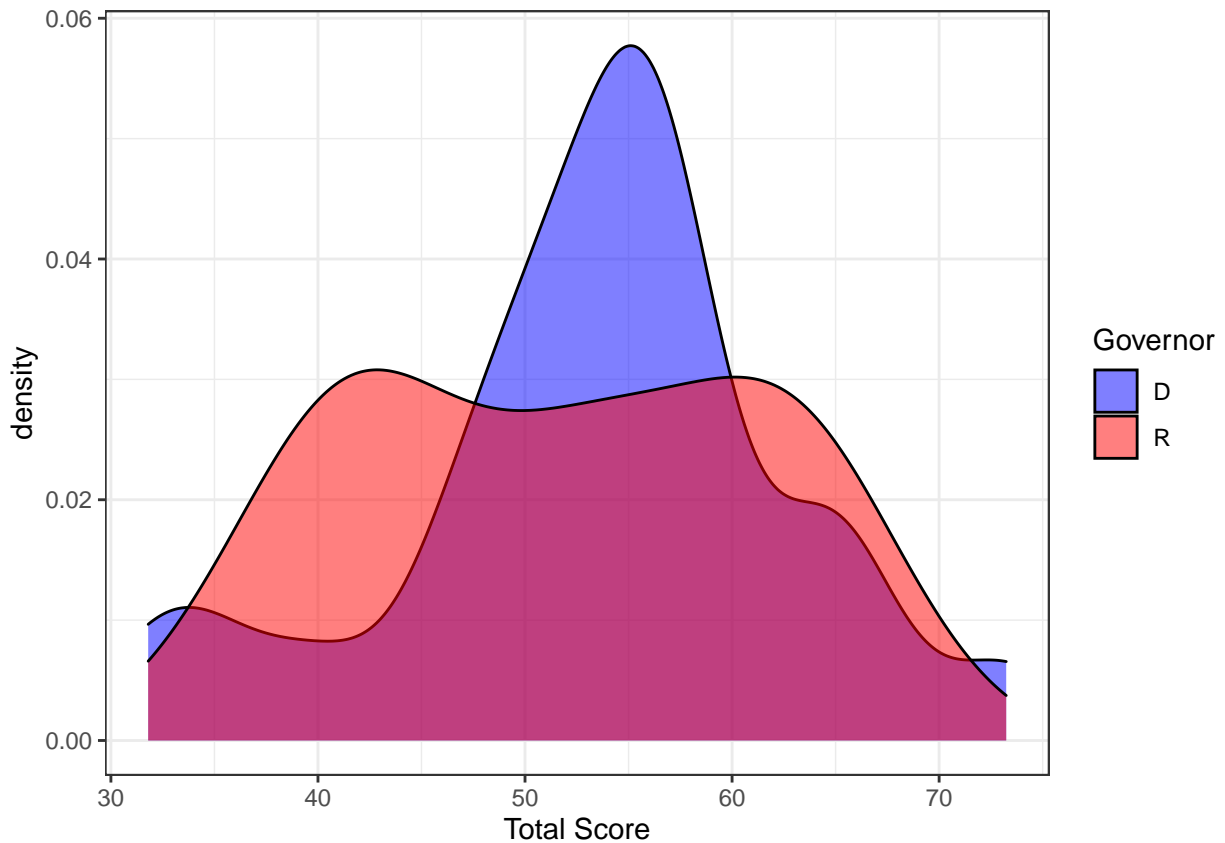
Predictor	B	SE	t	p
Intercept	55.11	2.136	25.80	<0.001
Republican Legislature	-5.07	2.702	-1.88	0.067
Democratic Legislature	11.35	6.754	1.68	0.099

Table 4: Legislative Party Control's Effect on Tax Burden: Tukey's HSD

Predictor	Comparison	Null	Estimate	Low	High	p
Party Control	R-D	0	-5.07	-11.610	1.47	0.157
Party Control	S-D	0	11.35	-4.994	27.70	0.223
Party Control	S-R	0	16.42	0.407	32.44	0.043

## Governors

The difference between states with Republican and Democratic governors is not statistically significant, though states with Democratic governors seem to vary less than those with Republican governors.



Welch Two Sample t-test

```
data: ranks$Total.Score by ranks$Governor
t = 0.65571, df = 47.092, p-value = 0.5152
alternative hypothesis: true difference in means between group D and group R is not equal to 0
95 percent confidence interval:
 -3.723303  7.324430
sample estimates:
mean in group D mean in group R
    53.49130     51.69074
```

## Overall Control

When we include the Governorship, control comprises 15 Democratic states, 24 Republican and 11 split between the two; however, overall party control isn't significantly correlated with WalletHub's service quality scores either.

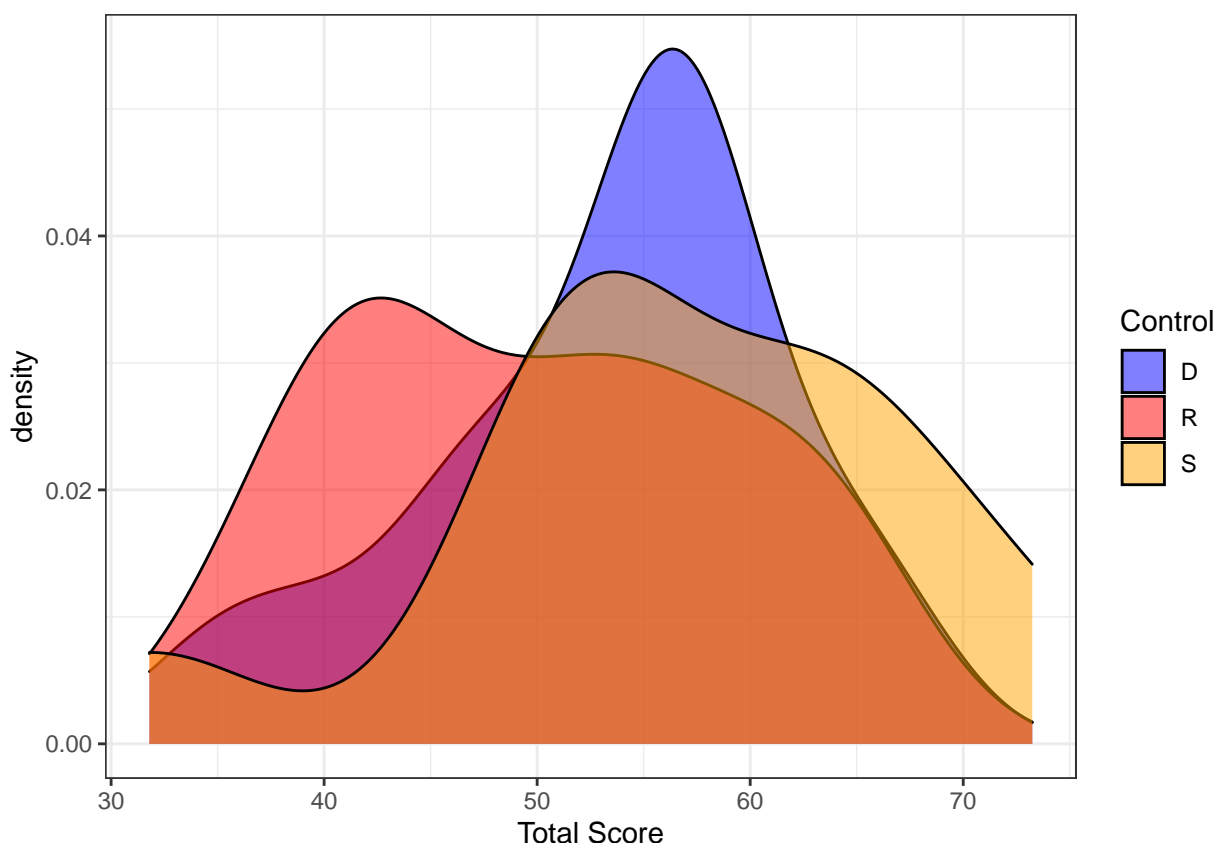


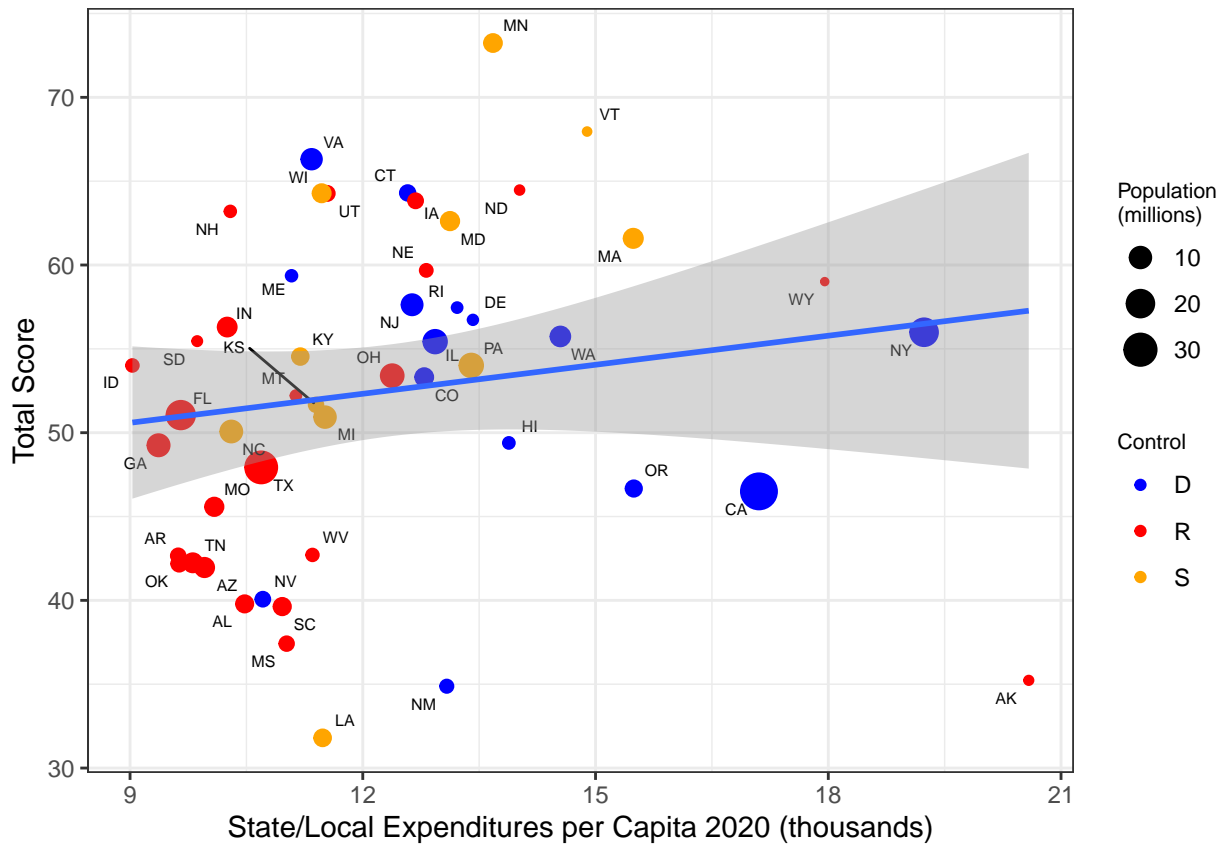
Table 5: Overall Party Control’s Effect on Quality Score: Linear Estimate

Predictor	DF	SumSq	MeanSq	Stat	p
Party Control	2	328.85	164.43	1.829	0.172
Residuals	47	4225.06	89.89	NA	NA

## Tax and Spend

Despite the problems with WalletHub’s analysis, the concept is intriguing, so we decided to look more closely at the relationship between revenue, spending and service quality. Although states collect money in different ways and amounts, wouldn’t their spending reveal more about their priorities and be a better predictor of service quality? What happens when we look at state and local spending in specific areas? Are there better ways to evaluate government service performance?

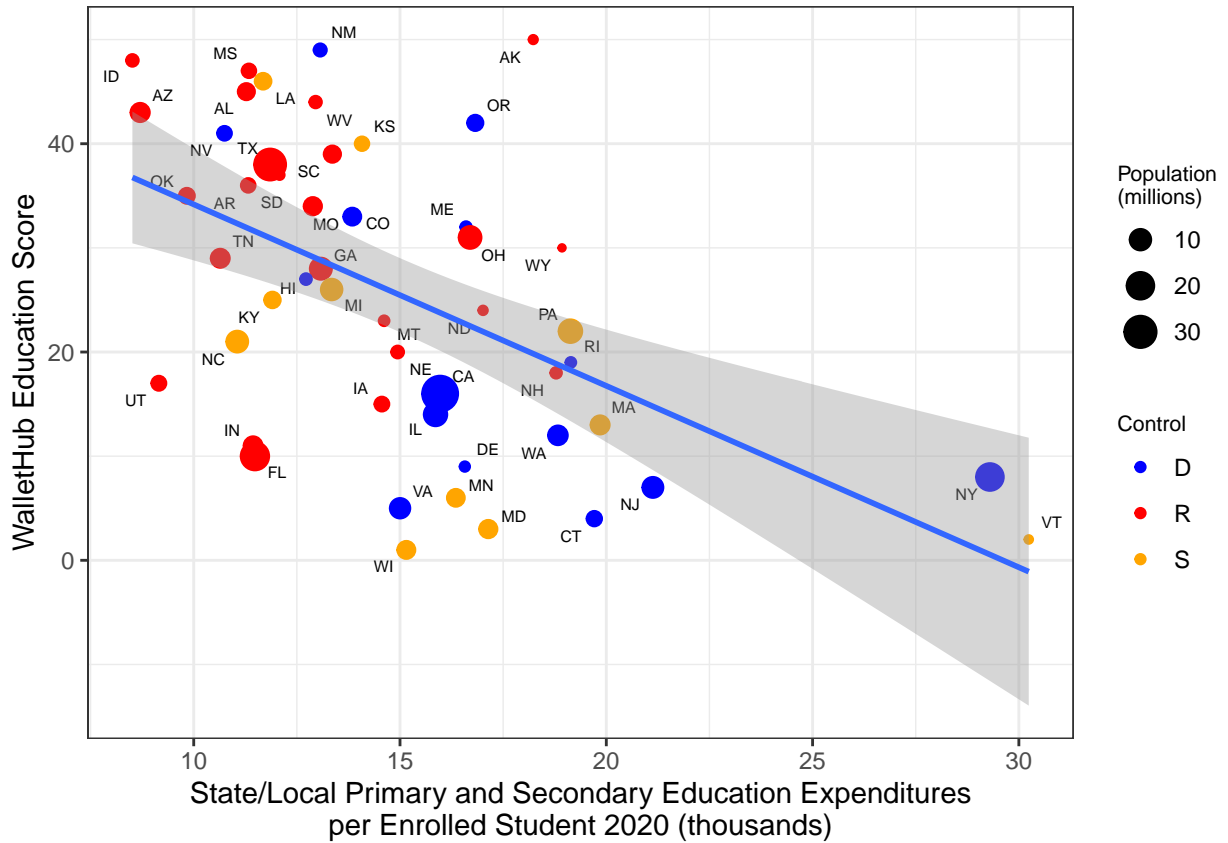
Rather than fixating on taxes, we instead looked at total state and local expenditures, including federal transfers, which should have a stronger relationship with government service quality. To do this, we included a dollar amount for [local and state revenue and expenditures](#) from the 2020 Census. But when we look at overall state and local expenditures per capita, there isn’t a statistically significant relationship with WalletHub’s quality score, which is surprising given what we found when we ran the same analysis with the tax burden.



## Education

More surprising is the strong negative correlation between spending on education and WalletHub's education score. As these are heavily weighted with two of WalletHub's other ranking indexes for [public schools](#) and [universities](#), it could be that this index of indexes is over-abstracted to the point it is no longer strongly connected to student outcomes.





This raises some questions about WalletHub’s education score methodology, since we can generate counter-examples. When we compare [8th grade math and reading scores from 2019](#), the most recent year available, we see a significant correlation between per student spending on primary and secondary education and test scores. CA and NY are notable exceptions.

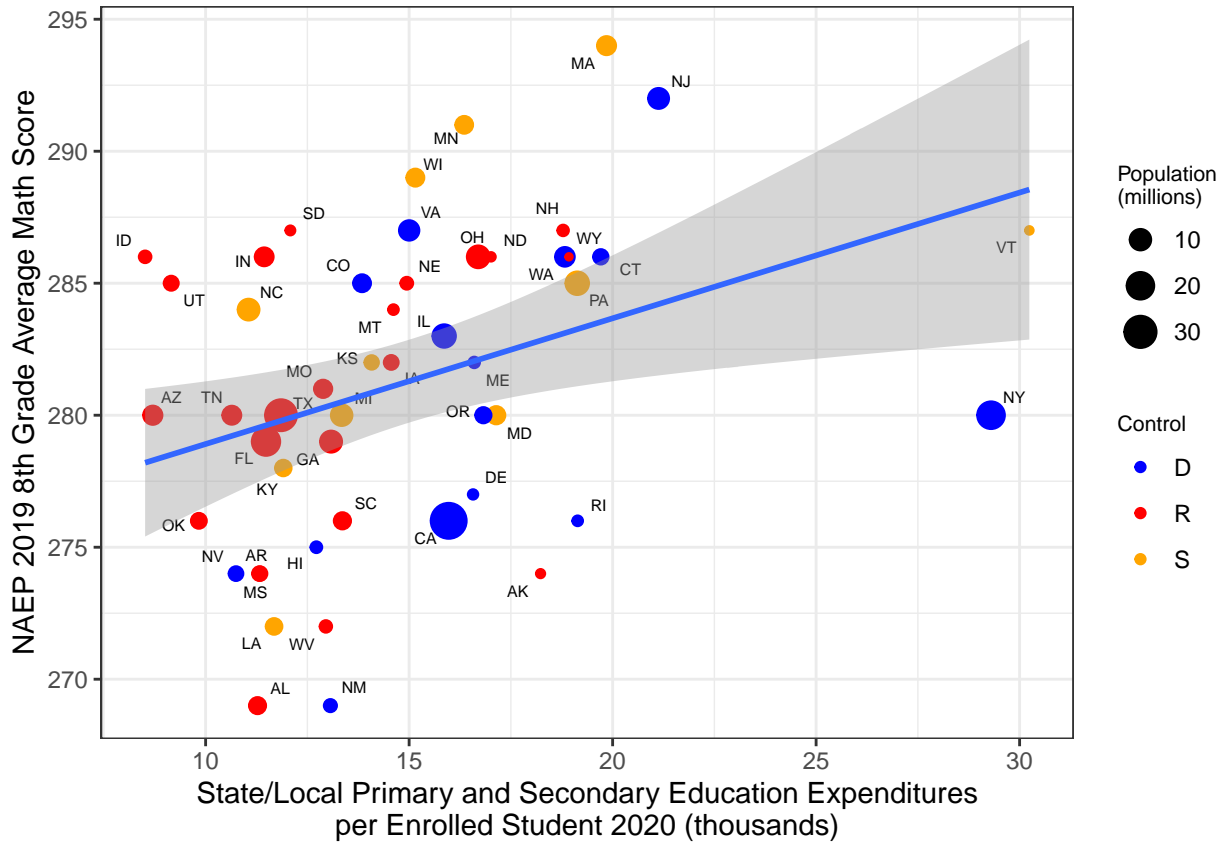


Table 6: Per Student Education Spend's Effect on NAEP Math Scores: Linear Estimate

Predictor	B	SE	t	p
Intercept	274.15	2.778	98.70	<0.001
Education Spend per Student	0.48	0.178	2.68	0.010

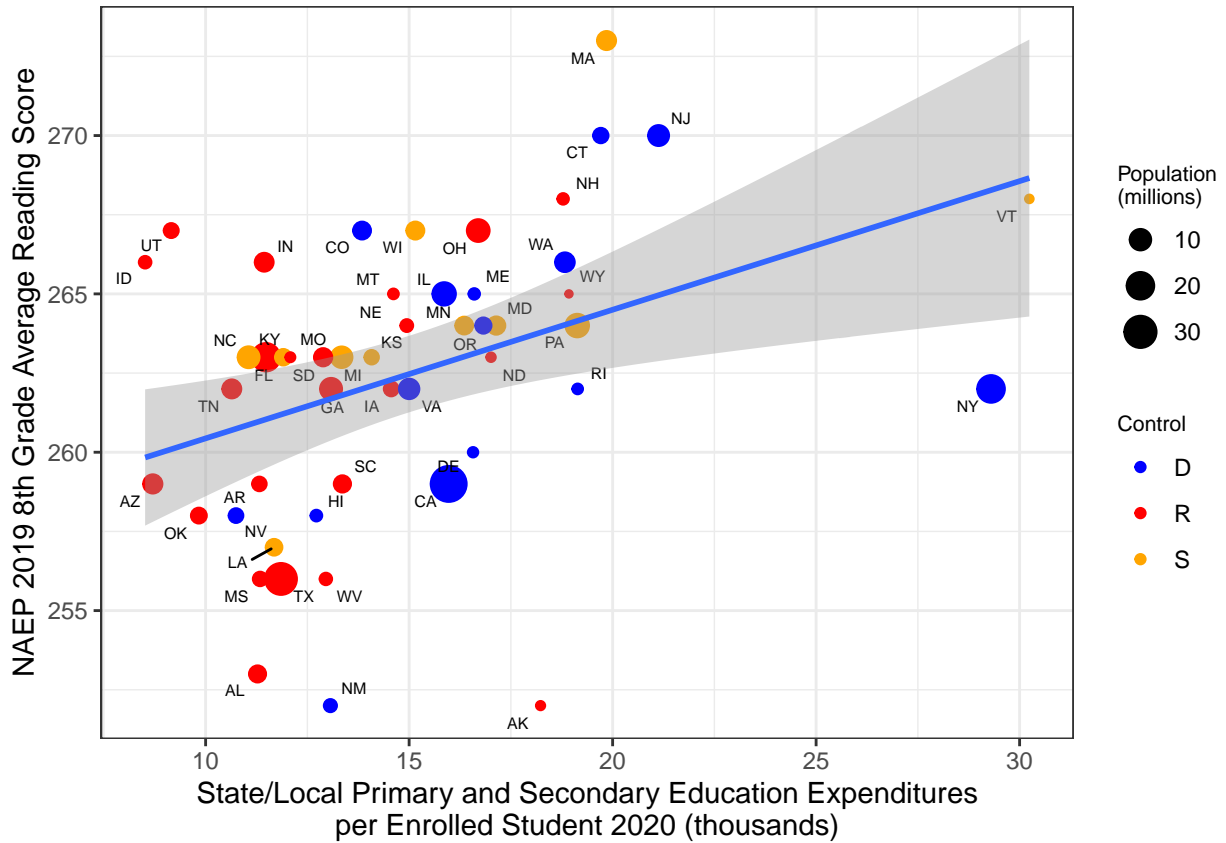


Table 7: Per Student Education Spend's Effect on NAEP Reading Scores: Linear Estimate

Predictor	B	SE	t	p
Intercept	256.38	2.137	119.94	<0.001
Education Spend per Student	0.41	0.137	2.97	0.005

## Health

Similarly, state and local per capita spending on health isn't significantly related to WalletHub's health score; however, we can generate counterexamples when we look at specific indicators like life expectancy, infant mortality and obesity.

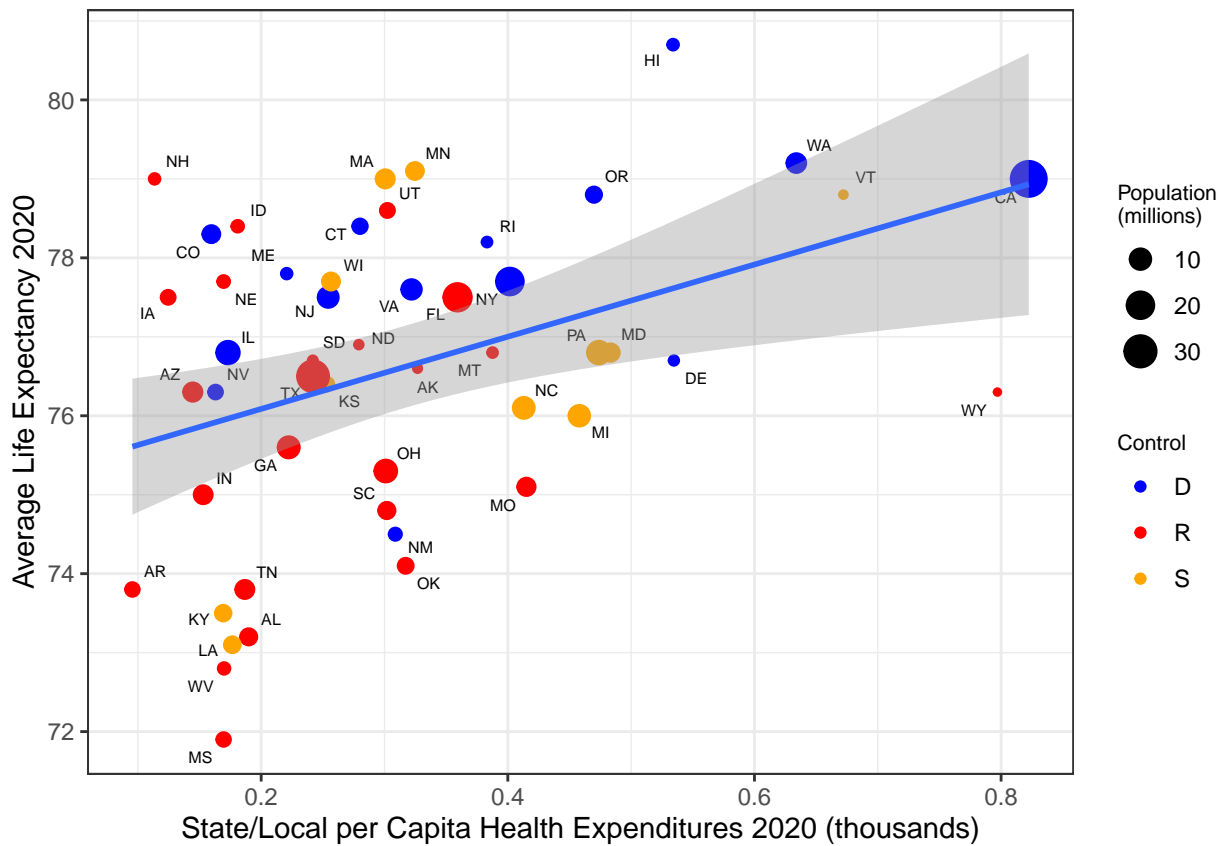
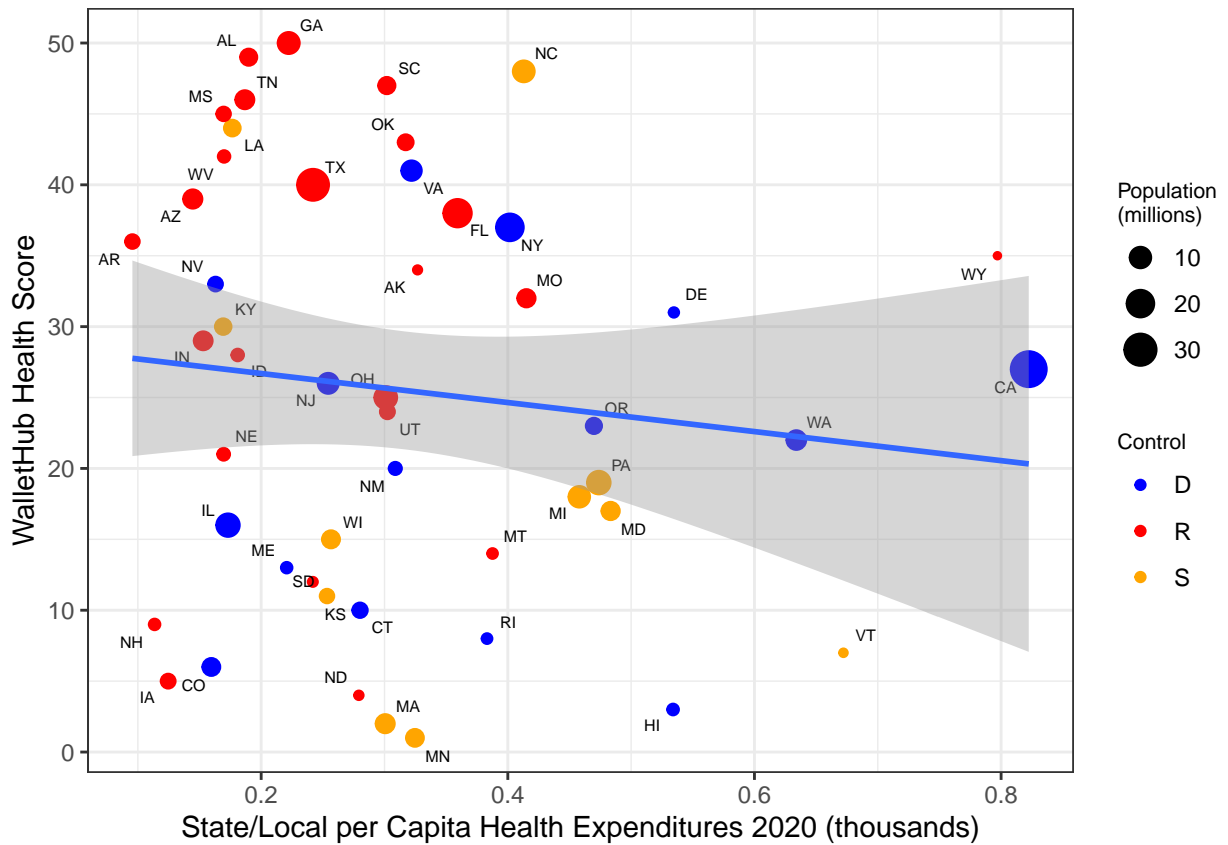


Table 8: Per Capita Health Spend's Effect on Life Expectancy:  
Linear Estimate

Predictor	B	SE	t	p
Intercept	75.17	0.554	135.78	<0.001
Health Spend per Capita	4.57	1.546	2.96	0.005

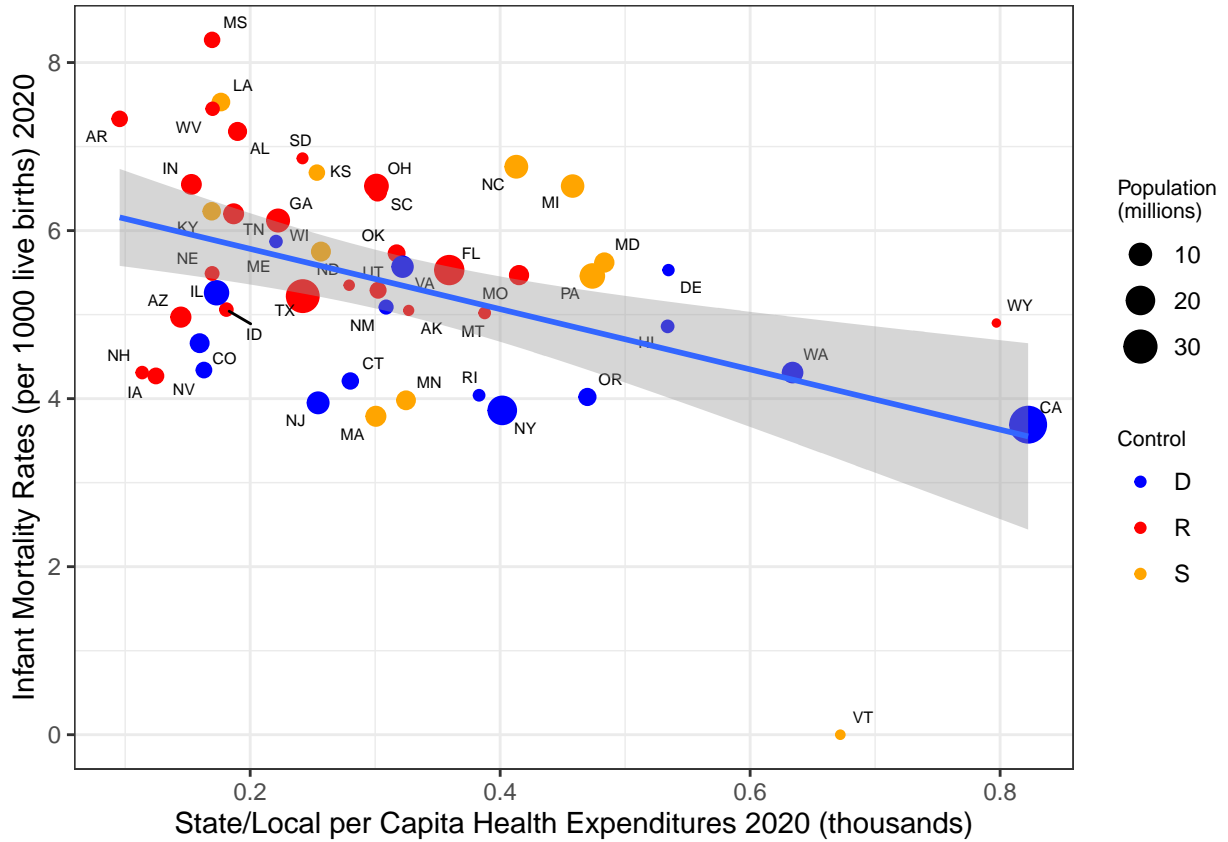


Table 9: Per Capita Health Spend's Effect on Infant Mortality:  
Linear Estimate

Predictor	B	SE	t	p
Intercept	6.50	0.371	17.53	<0.001
Health Spend per Capita	-3.59	1.035	-3.46	0.001

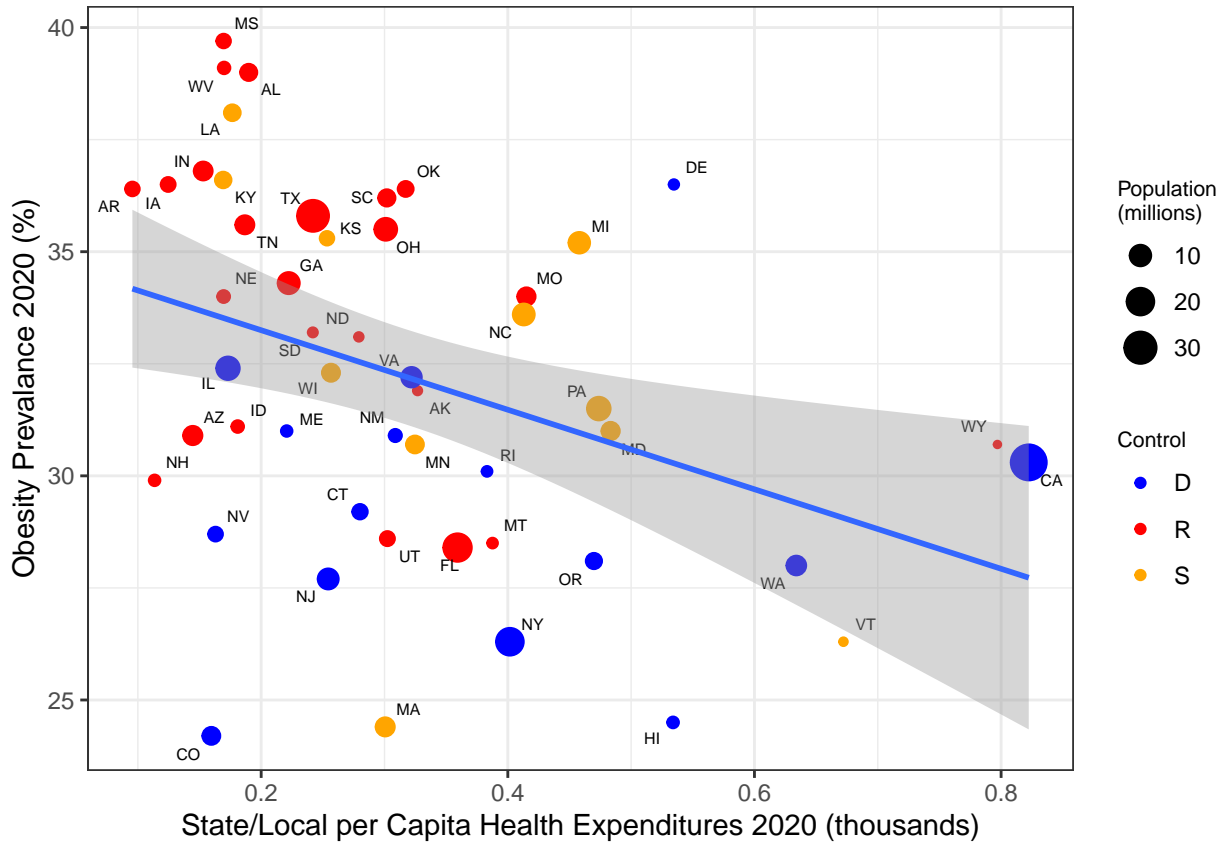


Table 10: Per Capita Health Spend's Effect on Obesity Rates:  
Linear Estimate

Predictor	B	SE	t	p
Intercept	35.02	1.131	30.95	<0.001
Health Spend per Capita	-8.87	3.160	-2.81	0.007

To be fair, life expectancy and infant mortality were included in the WalletHub health score, though again, two other WalletHub indexes, one on [Covid safety](#) and another on [Health Care Quality](#) were given significantly greater weight, along with estimates of costs which may not have been adjusted for income nor purchasing power. As both the WalletHub score indices as well as tax and spending amounts are point in time snapshots, they may not be indicative of longer term trends. For example, it is possible that in outlier states like CA and NY that spend heavily per capita, spending has only increased recently and outcomes reflect historical under funding. It is also possible that these correlations result from confounding variables not identified in this analysis.

## Conclusion

While WalletHub's analysis approached complicated questions about state tax policies and how these affect outcomes in key areas in a logical and reasonable way, on closer examination their measurement indices appear opaque and overly abstracted. It could well be that their methodologies are sound, and that the counterexamples presented here simply miss the forest for the trees; however, this is far from certain. The question raised is whether and how citizen outcomes can be connected to state revenues and expenditures, which is vital to voters and policymakers alike.

One thing we can conclude is that, at least in the one year snapshot we observe here, party control does not strongly influence outcomes for citizens. Perhaps this will change as polarization drives ever more radical and diverging choices between ‘Red’ and ‘Blue’ states, or perhaps the policy choices making headlines will not cause significant change in the basic functioning of transportation, health, education and other major elements of state and local infrastructure.

More in-depth examinations of specific policies and their outcomes would be useful, and the most important conclusion is that measurement matters. If WalletHub’s methodology is not perfect, it is an instructive and well-crafted first attempt from which to iterate. Is it reliable, for example, to base quality indices on other indices and retain accuracy and reliability to outcomes? Is there a generalized, single indicator that we can use to compare policy impact, like [this attempt](#) to apply the Genuine Progress Indicator to U.S. States, the [Human Development Index](#), [Measure of America](#) or OECD’s [Better Life Index](#)?

Or is it necessary to look at discrete state operations, their funding and small baskets of specific indicators over time? While this analysis illustrated some weaknesses of arbitrary composite indices, it is difficult to confidently identify specific indicators for outcomes that will be comparable across states given the huge variation in government spending and operational structures. What is more certain is that focusing on partisan rhetoric and power is the least constructive way to frame these discussions.