

Which is the Most Pro-Ron Paul State? Assessing the Determinants of Paul's Primary and Caucus Support

Jason Sorens

University at Buffalo

Department of Political Science

jsorens@buffalo.edu

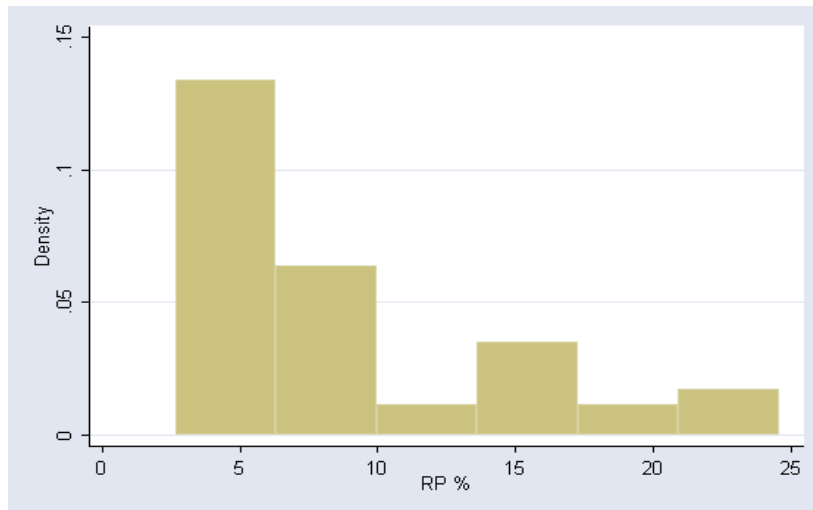
During the presidential primary season, many libertarians supported Ron Paul's candidacy. Indeed, Ron Paul's campaign ignited a grassroots libertarian movement dedicated to taking the Republican Party back to its supposed small-government roots. A number of enthusiastic Paul backers closely watched the election results and inferred from those numbers the strength of the libertarian movement in each state. Unfortunately, most of this seat-of-the-pants statistical analysis looked only at raw percentages and thus drew wildly incorrect conclusions. For instance, most Paul supporters were greatly disappointed when their candidate significantly underperformed the polls in New Hampshire, but some of them went further and castigated the "Free State" as unworthy of its reputation after Paul apparently did better in later caucuses. This paper remedies the flaws in prior, informal takes on Ron Paul's primary performances and tries to settle once and for all the question, Which state *really* backed Ron Paul the most?

To answer this question, I collected the Republican presidential primary and caucus results for all states for which data were available. Hawaii, Louisiana, West Virginia, and Wyoming had to be excluded, because they did not report Ron Paul's percentage of the vote. For Washington, which had both a caucus and a primary, only the primary results were included in the final estimations. The District of Columbia is included, but not other federal territories such as American Samoa. The goal of the paper is to explain Ron Paul's vote percentages in each state and D.C. as a function of turnout, number of candidates, caucus versus primary, and whether the contest occurred after John McCain had already won a majority of delegates, thus essentially clinching the nomination.

The histogram below shows the distribution of Ron Paul's state-level vote shares. Basically, what this figure shows is that the distribution of votes by state was highly skewed. Most states gave Ron Paul a very low vote share, but in some contests he did much better. Obviously, it is impossible for a vote percentage to go below zero, so it is not surprising to see so many results clustered at the low end, in states where perhaps the Paul campaign had very little organization.

Ron Paul's average share of the vote in the 47 states and D.C. was 8.6%. When caucus and primary states are separated, the average vote share in caucus states was 15.6%, while the average vote share in primary states was 6.9%. Thus, on average, Ron Paul did 8.7 percentage points better in caucuses than primaries. The reason for this large discrepancy is that caucuses require voters to be present at a single time in a particular place to cast their vote; voting is a much bigger commitment than simply going into a booth in your neighborhood at your convenience, which is how primaries and general elections work. Thus, candidates with more "diehard" supporters tend to do better in caucuses

Figure I: Distribution of Ron Paul Results



than primaries. (This is the same reason that Obama outperformed Clinton in caucuses.) Ron Paul had a more committed activist base than any other candidate and was thus able to turn out supporters for caucuses much better than the other candidates.

However, you can't simply subtract 8.7 from every caucus vote share to see how Paul would have done in a primary in that same state. Paul's worst caucus performance was 8.4%, in Colorado, but clearly a vote share of -0.3% is impossible. Because the distribution of vote shares is skewed, the relationship between the independent variables, such as caucus versus primary, and the dependent variable, vote shares, is closer to logarithmic than linear.

To figure out which variables predict Ron Paul's electoral performance, we therefore need to transform the raw vote shares through a natural logarithm before using a regression equation. In addition, the independent variables "turnout" and "number of candidates," which are described below, are skewed in the same way, because they cannot fall below zero. They are also logarithmically transformed. The regression equation thus looks like this:

$$\ln(\text{Paul \% of vote}) = a + b_1(\text{caucus}) + b_2(\ln(\text{turnout})) + b_3(\text{clinched}) + b_4(\ln(\text{number of candidates}))$$

The independent variables are as follows:

- "Caucus" – This variable is scored "1" if the state used a caucus election, "0" if the state used a primary. I expect this variable to be positively correlated with the log of Ron Paul vote percentages.
- "Turnout" – This variable measures voter turnout as total number of votes in the Republican presidential primary divided by total July 2007 population from the Census, times 100. This isn't quite the true definition of turnout, which is votes cast divided by eligible voting population, but figures for eligible voting population aren't available for the primaries, and these figures should not be biased so long as the percentage of the population that is eligible doesn't vary widely

across states (and it doesn't). I expect this variable to be negatively correlated with the log of Ron Paul vote percentages, because Ron Paul's more dedicated voters should be more likely to come out to vote than voters for other candidates.

- "Clinched" – This variable is scored "1" for elections after May 4, when John McCain clinched the Republican nomination. I expect this variable to be positively correlated with the log of Ron Paul vote percentages, because as the only active candidate, Ron Paul could expect to receive votes from many Republicans who disliked John McCain but did not necessarily support Ron Paul over other eliminated candidates no longer on the ballot. Also, again, Ron Paul's determined cadres might not have been very daunted by the fact that it was virtually impossible for him to be nominated at that point.
- "Number of candidates" – This variable simply measures the number of Republican candidates on the ballot. I expect this variable to correlate negatively with Ron Paul support, because the availability of more options allows protest voters or Paul-ambivalent libertarians to select another candidate, such as Fred Thompson in early contests or even "Vermin Supreme" in New Hampshire.

The election data were taken from the Wikipedia page, "Results of the 2008 Republican Party presidential primaries," which references each state's data to official sources.

Table I below shows the results of the multiple regression analysis.

Table I: Regression Results

Regression with robust standard errors

Number of obs = 47
 F(4, 42) = 25.98
 Prob > F = 0.0000
 R-squared = 0.6658
 Root MSE = .37562

	Coef.	Robust Std. Err.	T	P> t	[95% Conf. Interval]	
Ln(turnout)	-.1904951	.1474009	-1.29	0.203	-.487962	.1069719
Caucus	.8768918	.2004983	4.37	0.000	.4722699	1.281514
Ln(candidates)	-.1271726	.1488808	-0.85	0.398	-.4276263	.1732811
Clinched	.7330183	.2043432	3.59	0.001	.3206371	1.1454
Constant	2.185304	.3653621	5.98	0.000	1.447974	2.922635

The results show that a caucus increases Ron Paul's percentage of the vote by 0.88 "log points." Translating "log points" into percentage of the vote depends on what Ron Paul's percentage of the vote is otherwise expected to be. For instance, if the other independent variables predict that Ron Paul's percentage of the vote in a primary should be 5%, then changing the primary to a caucus raises Paul's predicted share of the vote to 12%. However, if Paul's vote is otherwise expected to be 8% in a primary, then changing the primary to a caucus raises Paul's predicted share of the vote to 19%. The boost that Ron Paul got in primaries after McCain clinched is almost as big: 0.73 log points. Turnout and number of candidates have smaller effects. Increasing turnout from 1%, as in the District of Columbia, to 10%, as in Missouri, decreases predicted Paul vote share by about a third of a log point. Increasing number of

candidates on the ballot from four, as in many Super Tuesday states, to 10, as in South Carolina, decreases predicted Paul vote share by only about one-ninth of a log point.

Once we control for all these factors, where did Ron Paul do best? One factor we have not been able to control for is the strength or efficacy of the Ron Paul organization. Presumably this is the “omitted variable” that explains most of the remaining variation across states. We can figure out where Ron Paul’s organization must have been strongest by looking at those states that significantly underperform or overperform the baseline model. The mechanics of this operation are as follows. First, we substitute our estimated coefficients into the regression equation, thus:

$$\ln(\text{Paul \% of vote}) = 2.185 + 0.877 * (\text{caucus}) - 0.19 * (\ln(\text{turnout})) + 0.733 * (\text{clinched}) - 0.127 * (\ln(\text{number of candidates}))$$

Then we substitute for each of the independent variables (caucus, turnout, clinched, candidates) the particular value for each state. The resulting $\ln(\text{Paul vote})$ figure is the *predicted value* of the log of Ron Paul’s vote share for each state. We can subtract this predicted value from the actual value, and the difference is the amount by which Ron Paul outperformed what we would expect, given primary or caucus format, turnout, timing (whether McCain had clinched or not), and number of candidates on the ballot, all factors outside Paul’s control. High values of (actual vote – predicted vote) indicate that Paul did better in those states than we would expect; low values mean he did worse. Therefore, high values presumably indicate that Paul had a better, more effective organization.

Table II ranks the states by the number of log points by which Paul outperformed his predicted performance in each state.

Table II: Paul’s Unexplained Performance by State

Overperformance	State
0.7993741	New Hampshire
0.7624688	Idaho
0.4739113	South Dakota
0.4645284	Washington
0.3769193	North Dakota
0.3725045	Pennsylvania
0.3609405	Michigan
0.3455322	Montana
0.296046	Oregon
0.2901108	Vermont
	District of
0.2458186	Columbia
0.2377231	Tennessee
0.2362665	Maryland
0.1911311	New York
0.1876662	Maine

0.1718122	California
0.1714756	New Mexico
0.1538725	Alaska
0.1459794	Nebraska
0.123376	Rhode Island
0.075878	Illinois
0.0525227	Missouri
0.0177805	Minnesota
0.0128961	Nevada
0.001461	Arkansas
-0.0221498	New Jersey
-0.0315268	Ohio
-0.0460235	Wisconsin
-0.0604255	Texas
-0.0773115	Virginia
-0.1129332	Arizona
-0.1456997	South Carolina
-0.1672701	Delaware
-0.1790915	Iowa
-0.231692	Oklahoma
-0.2629383	Connecticut
-0.2664576	Florida
-0.2991312	Kansas
-0.3194487	Indiana
-0.3745465	Georgia
-0.4054149	North Carolina
-0.4389548	Utah
-0.444923	Kentucky
-0.4810518	Alabama
-0.5319622	Massachusetts
-0.6164439	Colorado
-1.052599	Mississippi

So we can now answer the question: In which state did Ron Paul's organization do best? The answer is New Hampshire. New Hampshire was the most pro-Ron Paul state, although Idaho ran a close second. The other states are well behind. At the other end of the spectrum, Paul did really badly in Mississippi. We can only deduce that voters and activists in Mississippi strongly disliked or ignored Dr. Paul's antiwar, libertarian message. Colorado, Massachusetts, Alabama, Kentucky, Utah, and North Carolina were also pretty poor states for him.

New Hampshire of course is the home of the Free State Project (FSP), and the FSP has so far gotten about 300 or more libertarian activists to move to that state. If we assume that without the FSP,

New Hampshire would have been an average state for Ron Paul, neither underperforming nor overperforming our model, then we can figure out how many extra votes these “Free Stater” activists obtained for Ron Paul. Given the highest turnout and number of candidates in the nation, the fact that the race was in the early stages, and the primary format, the model predicts that Ron Paul should have gotten 3.44% or 8,234 votes in New Hampshire’s primary, but he actually got 7.65%, 18,308 votes. Since there were about 500 Free Staters in New Hampshire at the time (300 early movers, 200 FSP participants who lived in New Hampshire before the state was chosen as the FSP’s destination), each Free Stater obtained 20.1 (10,073/500) extra votes for Ron Paul.

Of course, this is a high-end estimate, since Operation Live Free or Die, which brought in activists from around the country to New Hampshire for a few weeks before the primary, certainly accounts for some of those extra votes. But if the FSP were not in New Hampshire, would Operation Live Free or Die have attracted as many activists or even focused on New Hampshire at all, as opposed to Iowa or South Carolina?

The ultimate goal of the FSP is to attract 20,000 libertarian activists to New Hampshire. If each activist generated 20 votes for libertarian candidates, that would mean 400,000 votes for libertarian candidates at each election. In the 2004 presidential election, only 676,000 votes were cast in New Hampshire. Clearly, we have to assume some diminishing marginal returns. However, even if only 10,000 of those activists are as effective on average as the activists already in New Hampshire, then libertarian candidates could get 200,000 votes statewide consistently. That would perhaps not be enough to hand New Hampshire’s electoral votes to a Libertarian Party candidate in a three-way race, but it would be enough to dominate all statewide primary elections for either major party. Only 239,000 votes were cast in the 2008 Republican presidential primary, and fewer than 60,000 votes were cast in the last high-profile statewide gubernatorial primary, the 2004 Democratic contest. Thus, even 1,500 activists (30,000/20) might be enough to win pretty much every statewide primary contest in either major party.

Should Ron Paul fans be disappointed in his primary election results in New Hampshire? Quite the contrary! Controlling for turnout and other factors, New Hampshire was his best state.