Does Challenger Quality Matter?

Candidates and Donors in the 2006 Gubernatorial Elections

Adam R. Brown

Dept of Political Science

Brigham Young University

http://adambrown.info/

Prepared for delivery at the 2008 Annual Meeting of the American Political Science Association,

August 28-31, 2008.

Abstract: In a recent paper I showed that potential challengers and their financial backers behaved strategically when deciding whether to challenge incumbent governors in 2006; governors who appeared vulnerable in late 2006 attracted the most experienced, best-funded challengers (Brown 2008). In this paper, I ask whether challenger strength even matters by examining two outcomes of interest: Changes in the incumbent's popularity over the course of the campaign, and the incumbent's performance on election day. Challenger strength turns out to matter only weakly, with a stronger effect on election results than on incumbent popularity. This difference suggests that if challenger strength matters, it does so partly because inexperienced challengers do not provide voters with a real choice on election day—experienced challengers are not necessarily better campaigners than inexperienced ones.

The twenty-six governors seeking reelection in 2006 faced every sort of challenger imaginable. A few lucky incumbents faced challengers with no political experience and minimal campaign resources. Less fortunate incumbents faced richly experienced or well-funded challengers capable of running a serious campaign. Not surprisingly, incumbent governors performed better on election day if they had one of the weaker challengers. Nebraska's popular Dave Heineman won 75 percent of the two-party vote against his politically inexperienced, poorly funded challenger; Maryland's Ehrlich won only 46.7 percent against the exceptionally well-financed mayor of his state's largest city. More broadly, a simple regression of the incumbent's two-party vote share on the challenger's experience and spending shows that these two variables alone explain 51 percent of the variance in the 2006 gubernatorial election results.[1]

This strong relationship is unsurprising; we already know that stronger challengers fare better in House (Jacobson 1989) and Senate (Lublin 1994) elections. Nevertheless, the fact that strong gubernatorial challengers outperform weak ones does not mean that challenger strength matters in any way. We know from the Congressional literature that potential candidates and their financial backers behave strategically when deciding whether to challenge the incumbent (Jacobson 2004). Recent research has confirmed this pattern in the gubernatorial context (Brown 2007; Brown 2008; Brown and Jacobson 2008; Squire 1992; but see Leal 2006). Since experienced, well-financed candidates challenge only the most vulnerable incumbents, it could be that strong challengers outperform weak challengers on election day merely because the incumbents they oppose were weaker to begin with. Returning to the two examples cited earlier, Nebraska's Heineman won 75 percent of the two-party vote—but his approval rating was already at 71.4 percent a full year before the election, long before the challenger's identity was even known. Likewise, Maryland's Ehrlich won only 46.7 percent of the two-party vote on election

day, but that was not a far departure from his 51.9 percent approval a year earlier.

Unless strong challengers have an independent effect apart from the conditions that prompted their candidacy, then the fact that potential challengers and donors behave strategically when deciding whether to take on the incumbent becomes a meaningless tidbit of only academic interest. In this paper, I take a detailed look at the 2006 gubernatorial elections to assess the real effects of challenger strength, if any, in that context. I examine the effects of challenger strength on two outcomes of interest: The incumbent's popularity over the course of the campaign, and the incumbent's share of the two-party vote on election day.

From a theoretical standpoint, there is no reason to expect challenger strength to have the same effect on approval ratings as on election results. Elections, after all, are different from approval surveys. On election day, voters evaluate the incumbent relative to the challenger; if the challenger is politically experienced and well-known then it seems reasonable that voters desiring change might be willing to put their faith in her, but if the challenger is a political neophyte then even those who intensely dislike the incumbent may hesitate to place the challenger in office. In approval polls, by contrast, respondents evaluate the incumbent in isolation, not in relation to the challenger. Challenger quality may matter indirectly in approval polls, at least to the extent that political experience and campaign spending help challengers formulate more effective attacks on the incumbent, but it might not.

More precisely, any effects of challenger quality on approval ratings or election results may arise from two separate sources. First, it might be that strong challengers run better campaigns; second, it might be that campaigns do not matter at all, and strong challengers outperform weak challengers only because voters are unwilling to abandon an incumbent, no matter how unpopular, for a political rookie. When analyzing election results alone, we observe

the simultaneous effects of both these mechanisms and cannot disentangle them statistically. But when analyzing approval ratings over the course of the campaign, we observe only the first mechanism. To the extent that challenger quality has different effects on approval ratings than on vote shares, then, we glean insights about the relative importance of these two causal mechanisms.

After taking account of the incumbent's initial vulnerability, several intriguing findings emerge in the empirical analysis below:

- Political experience benefits challengers (slightly) on election day, but it does not benefit them during the campaign—that is, the challenger's experience does not influence the incumbent's approval rating even though it does hurt the incumbent's vote share. This finding implies that politically experienced challengers outperform inexperienced challengers mostly because they present voters with a real choice on election day—not because they are better campaigners.
- The challenger's campaign spending, in contrast, has roughly the same effect on approval ratings as on vote shares, with some small but theoretically suggestive differences that I discuss below. In this respect, then, challenger quality does matter; well-funded challengers can strengthen the tide against an incumbent.
- Surprisingly, and in stark contrast to findings in the Congressional context, the incumbent's campaign spending also matters—and even more strongly than the challenger's spending does. Since incumbents typically outspend challengers, the net effect of challenger and incumbent spending almost always works in the incumbent's favor. This reality creates a strong incentive for the best potential candidates to behave strategically when choosing whether to challenge the incumbent; otherwise, even the

strongest potential challengers are almost certain to lose.

I now turn to my empirical analysis. Following a few comments about data and measurement, I first present my analysis of the incumbent's approval ratings over the course of the campaign, followed by a look at the incumbent's performance on election day. In each section, I begin by setting up a baseline model that predicts the outcome using only those variables shown elsewhere to predict challenger strength (Brown 2008), discussed below. I then add my two indicators of challenger strength (i.e., experience and finance) into this baseline model, separately and then together; if these variables do not improve on the baseline regression, then we must conclude that challenger strength has no independent effect and that strong challengers outperform weak ones merely because they exploit favorable conditions.

Data and Methods

Although gubernatorial approval data has been collected occasionally since the advent of polling, polls were commissioned only sporadically until 2005. From May 2005 through November 2006, SurveyUSA, a firm whose main clients work in journalism, collected monthly gubernatorial approval ratings in every state. Although SurveyUSA does not make individual responses available, it does provide approval ratings in the aggregate and by partisan subgroup. The reliability and validity of SurveyUSA's data has been discussed thoroughly elsewhere (Brown 2007; Brown 2008; Jacobson 2006). Thirty-six states held gubernatorial elections in 2006. I examine the twenty-six in which incumbents sought reelection, since the traditional strategic emergence theory makes limited predictions about open seats.

I measure each governor's initial vulnerability using averaged approval ratings from May through December 2005. Although this decision was arbitrary, other reasonable periods produce

comparable results. As reported in Brown (2008), these averages serve as remarkably good predictors of the eventual challenger's strength, whether measured as experience or as spending. In the Congressional context, researchers are accustomed to using a variety of analytical proxies in their efforts to estimate incumbent vulnerability, such as the national partisan trend (Jacobson 1989), the district's partisan tendencies (Bond et al. 1997; Westlye 1991), the incumbent's ideology and policymaking behavior (Bond et al. 1985), and the size of the incumbent's financial reserves, or "war chest" (Goodliffe 2001; Goodliffe 2007). SurveyUSA's direct measurements of gubernatorial popularity render these indirect proxies unnecessary. For good measure, though, I also include a partisan dummy and a measure of George W. Bush's state-level approval where appropriate as additional indicators of the incumbent's vulnerability, since these national variables have been shown to also have a weak influence on challenger strength (Brown 2008).

Challenger strength has two components, spending and experience. I measure challenger spending using contributions data collected by the National Institute on Money in State Politics, supplemented with information from Vermont's Secretary of State. My measure of challenger experience is derived from one of the measures used in Brown (2008), a four-category ordinal measure that was, for the most part, a monotonic transformation of the highest percentage of the state that the challenger had previously represented, with minor adjustments for legislative leadership and celebrity status. Before using this ordinal measure as a right-hand variable, I dummy it out; to preserve two statistical degrees of freedom, I first collapse the four categories into two. Challengers classified in either of the original variable's top two categories are here identified as "experienced."[2] For the most part, this means that challengers who had previously held statewide office, legislative leadership positions, or Congressional seats are experienced, while legislative backbenchers, less prominent officeholders, and newcomers are coded as

inexperienced.

Effects of Challenger Strength on the Incumbent's Popularity

To assess the effects of challenger strength on the incumbent's approval, I measure approval as the one-step change in each governor's approval rating between May-December 2005 and September-October 2006.[3] By differencing, I incorporate the incumbent's original vulnerability into the dependent variable, eliminating the need to control for it on the right-hand side. As such, the baseline model shown below controls only for the incumbent's partisanship,[4] which leaves much to be explained. On average, Democratic governors enjoyed a 4.26 percentage point rise in approval during the campaign, while Republicans saw a much smaller rise (see Table 1).

[Table 1 about here]

Does Experience (alone) Help?

The challenger's political experience does not appear to hurt the incumbent's popularity over the course of the campaign; unexpectedly, the challenger's prior experience actually had the opposite effect in 2006. As shown in Table 2, incumbents with experienced challengers enjoyed a five-point rise in their approval ratings over the course of the campaign relative to incumbents with inexperienced challengers. This odd finding weakens under some specifications of the model, but no specification yields a significantly negative estimate—although several specifications return a significantly positive estimate like the one shown below.[5] This puzzling pattern persists even when Republicans and Democrats are examined separately, although it is somewhat stronger among Democratic challengers to Republican incumbents.

[Table 2 about here]

This unexpected result arises as an artifact of some reversion to the mean that occurred between late 2005 and October 2006. Figure 1 makes this pattern apparent, with changes in approval plotted against late 2005 approval. (To keep the figure readable, the figure displays data only for Republican incumbents.) Reversion to the mean is a purely statistically phenomenon; other things being equal, those at the extremes on any variable in one period are likely to move toward the mean in the next. If no reversion to the mean had occurred, then the dots in this figure would form no pattern. Instead, we observe a clear negative relationship between late 2005 approval and changes in approval.

[Figure 1 about here]

Three of the most popular Republicans (in Connecticut, South Dakota, and Vermont) experienced the largest drops in approval; at the other extreme, the least popular governor (in California) experienced the largest rise. Due to their early popularity, the governors of Connecticut, Vermont, and South Dakota had attracted weak challengers; due to his early unpopularity, the governor of California had attracted a strong one. As such, challenger quality is unfortunately correlated with this statistical phenomenon[6]; these four influential observations bias the analysis toward finding a counterintuitive result.

A common solution to this statistical problem would be to introduce a lagged approval measure, but with the present sample size even this tactic does not change the result. Given this statistical bias, then, we cannot make any conclusions about the effects of challenger experience on approval ratings. While it is possible that challenger experience genuinely did have the counterintuitive effect reported in Table 2, it seems far more plausible that challenger experience simply did not affect incumbent approval ratings at all—particularly in light of additional findings presented later in this paper.

Incidentally, this statistical bias also works against most of the other analyses in this paper. When challenger strength is found to matter, then, it does so despite a bias to the contrary—reinforcing confidence in the results. If anything, the true effects of challenger strength, whether proxied as experience or spending, are stronger than reported below.

Does Money (alone) Help?

In contrast to challenger experience, challenger spending did have a clear, significant effect on the incumbent's popularity over the course of the 2006 campaign (see Table 3). Of course, analyzing the effects of campaign contributions is a tricky exercise. For one thing, it is rarely clear how to compare spending across states with unequal populations. I use the same measure developed in Brown and Jacobson (2008)—that is, the logged dollar amount raised by each candidate.[7] The estimated effects of spending change little when contributions by challengers to their own campaigns are removed from the spending variable. While it is true that the fit and estimated coefficients rise slightly when challenger self-finance is included, the aggregate differences between the two models are relatively small. For purposes of this section, I focus my discussion on the first model.

[Table 3 about here]

In these estimates, observe that challenger spending has a substantial, statistically significant estimated coefficient. This estimate suggests that strong challengers can successfully lower the incumbent's approval rating over the course of the campaign. Though technically true, however, there is a difficulty with this conclusion: The incumbent's spending has a slightly stronger estimated effect than the challenger's—in the opposite direction. Thus, if the incumbent's spending rises to counter the challenger's, incumbents might fare much better than the challenger's spending alone would lead us to expect. And in 2006, incumbent spending did

rise to match challenger spending; at the margin, every one percent increase in challenger spending led to a 0.6 percent increase in incumbent spending, an equation that explains 57% of the variance in incumbent spending.[8]

Challenger spending does matter, then, but incumbent spending matters just as much, if not more. Based on the coefficients reported in the first model above, the challenger would need to raise roughly 70% more money than the incumbent in order for the negative effect of the challenger's spending to outweigh the positive effect of the incumbent's spending. As it happens, only five challengers managed to outspend the incumbent in 2006; of these, only two raised the requisite 70% more than the incumbent.[9] Only in these two states would we expect to see a fall in gubernatorial approval during the campaign. On average, though, the twenty-six governors seeking reelection witnessed an average rise—not a fall—in their approval ratings of roughly 2.5 points over the course of the campaign.

This pattern of incumbent and challenger spending diverges sharply from what has been observed in Congressional elections. In that context, as in the gubernatorial context, incumbents spend reactively, raising only enough money to defeat their challengers. In Congressional elections, however, this reactive relationship between challenger and incumbent spending leads to a counterintuitive empirical result: The more the incumbent spends, the more likely he is to lose (Jacobson 2004). A simple explanation underlies this odd finding; incumbent spending does not affect Congressional election results, and since incumbents spend only when they feel threatened, their spending indicates their fears of impending loss. But as the results above show, gubernatorial elections differ from Congressional elections in that *both* the challenger's and the incumbent's spending matter.

Given the competing effects of challenger and incumbent spending in gubernatorial

elections, measuring each variable separately may not be the most straightforward way to assess the real-world effects of challenger spending on election outcomes. What matters on election day is not the raw amount spent by the challenger, but whether the challenger managed to outspend the incumbent. Although it may be statistically less precise, we gain a clearer understanding of the substantive effects of challenger spending when it is measured as a percentage of the two-candidate spending total.[10] On average, challengers were overwhelmingly outspent by incumbents; the typical challenger spent only 34.4% of the two-candidate total. The weakest challenger was lopsidedly outspent, spending a meager 4.4% of the total; the strongest challenger spent 68% of the total, roughly twice as much as his opponent.

As shown in Table 4, re-estimating the previous regression using this new variable produces slightly diminished fit but essentially the same coefficient on the partisanship dummy. However, the marginal effect of challenger spending becomes much easier to interpret under this new specification. For every percentage point increase in the challenger's spending as a share of the total, the incumbent's expected change in approval ratings falls by 0.18 percentage points.

[Table 4 about here]

Against Democratic incumbents, the spending variable alone explains over half the variance in the outcome. Against Republican incumbents, the fit is far poorer and the coefficient is statistically insignificant. Moreover, the estimated effect of spending is noticeably larger against Democratic incumbents, although neither estimated effect departs much from the estimate in the aggregate model. The reversion to the mean discussed earlier weakens these results slightly; introducing an approval lag causes the spending coefficient to increase to -0.16 ($p=0.10$, or 0.05 one-tailed) for Republican incumbents and -0.23 ($p=0.01$) for Democratic incumbents. Nevertheless, the substantive conclusions remain essentially the same even with this

change. It appears, then, that spending benefits challengers of both parties, with a slightly stronger effect against Democratic incumbents. I discuss this partisan difference further in the next section.

These analyses use aggregate approval data. Similar results obtain when measuring gubernatorial and presidential approval only among members of the governor's opposition. By contrast, the results deteriorate noticeably when approval is measured among independents or members of the incumbent's party. This pattern is consistent with Brown (2008), which identified out-party approval (measured in late 2005) as the best predictor of challenger strength. However, this finding does not mean that only members of the governor's opposition are receptive to campaign messages. In fact, Table 5 shows that in-party approval moved even more than out-party approval did over the course of the campaign. When evaluating Republican governors, in-party respondents showed four percentage points more change during the campaign than out-party respondents; when evaluating Democratic governors, the difference was almost six percentage points.[11]

[Table 5 about here]

However, this rise in approval signifies a rally by in-party respondents around their incumbent—not a response to the challenger's campaign. As discussed in Brown and Jacobson (2008) and later in this article, campaigns provide opportunities for incumbent governors to redeem themselves in the public's mind; in-party respondents are particularly receptive to positive information about the governor, leading to the large rises in approval among this group. The task for challengers is to raise enough money to counter the incumbent's positive messages, but the incumbent's partisans appear to pay far less attention than the challenger's partisans to the challenger's efforts. Among low-spending challengers to Republicans—that is, those who

raised less than 30 percent of the spending total—in-party gubernatorial approval rose 4.8 points while out-party approval rose 2.1 points. Among high-spending challengers, approval rose just as much among in-party respondents (slightly more, actually), but it hardly rose at all among out-party respondents.

Comparing these two differences shows that the effect of challenger spending was 2.4 percentage points stronger among out-party respondents than among in-party respondents. Where the governor was a Democrat, the effect was 3.6 points stronger among out-party respondents. For this reason, challenger spending has the strongest effects when out-party data is used. Although voters of all stripes reassessed the candidates during the 2006 campaign, out-party respondents appeared far more receptive to the challenger's appeals.

Do Experience and Money (combined) Help?

Combining the experience and spending analyses yields no new insights; their effects are additive, not interactive. Additionally, neither experience nor spending interacts significantly with the incumbent's late 2005 approval ratings. With such a small sample size, interactions are difficult to test, of course; adding additional variables drives the standard errors rapidly skyward. Nevertheless, whether analyzed in the aggregate or by partisan subgroup, the patterns reported above persist. When estimating the change in each incumbent's approval ratings over the course of the campaign, spending matters but experience does not; the effect of spending appears to have been strongest among Republican challengers to Democratic incumbents. I now turn to the effect of challenger strength on incumbent vote shares.

Effects of Challenger Strength on the Incumbent's Vote Share

If high-quality gubernatorial challengers in 2006 merely took advantage of favorable

conditions rather than strengthening them, then we would be unable to improve on the baseline models in Table 6. These models use the same right-hand variables that predict challenger quality in the first place (as discussed above), namely, the incumbent's popularity in late 2005, Bush's state-level popularity during the same period, and a partisanship dummy. As such, these variables summarize the "wave" that the challenger might ride against the incumbent. Both models in Table 6 use the incumbent's share of the two-party vote as the dependent variable. The first model uses aggregate approval measures; the second uses approval among the governor's opposition.

[Table 6 about here]

The three baseline variables are statistically significant in both models, at least for a one-tailed test. And either model accounts for a large amount of the variance in election results—between 55 and 59%. Even before a challenger enters the race, then, the incumbent's initial vulnerability has largely determined his eventual share of the two-party vote. If so much of the election result is determined before the challenger even decides whether to run, then we have a high statistical hurdle to overcome in order to demonstrate that challenger strength makes any additional contribution to these trends.

Although the two models return similar results, the out-party model outperforms the aggregate model slightly; not only is the fit slightly better with out-party data, but Bayesian model selection supports the out-party model as well.[12] By contrast, the results are much worse when estimated with in-party approval data, consistent with the discussion above.[13] As such, I use out-party data for the remainder of this analysis, although this decision is not consequential; similar results obtain when using aggregate approval data.[14]

Does Experience (alone) Help?

In 2006, challengers varied dramatically in their prior political experience. Several challengers had neither political experience nor widespread name recognition. At the other extreme, several challengers had previously held federal or statewide offices. In Brown (2008), I introduced two different measures of challenger experience: The highest percentage of the state's population that the challenger had ever represented in elected office (logged) and a more qualitative ordinal measure that adjusted for legislative leadership and celebrity status. As it turns out, however, neither of these variables adds anything to the baseline model in Table 6 when it comes to predicting the election result; neither has a statistically significant coefficient, and neither improves the fit. Dummying out the ordinal measure raises the R-squared estimate, but only because doing so introduces three new right-hand variables; the adjusted R-squared moves little (not shown).

Unexpectedly, these non-findings arise from a curious partisan interaction. In 2006, only Democratic challengers to Republican incumbents benefited systematically (but weakly) from prior political experience; Republican challengers did not. The following two tables split the sample; the first table shows a series of models predicting the vote shares for Republican incumbents, and the next table looks at Democratic incumbents.

In Table 7, OLS 1 displays the baseline model using data only for Republican incumbents. Like the aggregate baseline model, OLS 1 explains 59% of the variance in vote shares. In OLS 2, I insert the dummy measure of challenger experience discussed above. The model predicts that an experienced challenger can reduce the incumbent's vote share by 8 points on election day, a large effect; this estimate has one-tailed significance in the expected direction. In addition, inserting this variable improves the fit dramatically. The substantial rise from 0.59 to

0.73 in the R^2 (from 0.51 to 0.64 in adjusted R^2) is not a fluke; Bayesian model selection also gives positive support for OLS 2 over OLS 1.

[Table 7 about here]

Of course, these conclusions require considerable caution. First, splitting the sample to analyze Republicans and Democrats separately reduces the number of observations in each model from 26 to only 13, a very small number for this sort of analysis. Second, the effects discussed above are somewhat contingent on measurement decisions; when experience is measured as the logged percentage of the state that the challenger had previously represented, the estimated effect of experience is not statistically significant. All the same, the substantial difference between OLS 1 and OLS 2 is certainly suggestive of what we might find in a larger-N multi-year study.

By contrast, Republican challengers to Democratic incumbents did not seem to gain much from their political experience, as shown in Table 8. Once again, OLS 1 sets up the baseline model, which in this case explains a whopping 74% of the variance in the Democratic incumbents' vote shares. Adding the experience variable (OLS 2) does nothing whatsoever to improve on this high baseline. Not only does the fit not improve, but none of the other coefficients changes substantially. This finding repeats itself for several specifications of challenger experience. At least in 2006, Republican challengers gained nothing at all from their prior political experience.

[Table 8 about here]

It appears, then, that Democratic challengers benefited somewhat from their political experience while Republicans did not. This partisan interaction might be real, but it might also be an artifact of the electoral context in 2006. As discussed already, experienced challengers

tended to run only against vulnerable governors. A major component of vulnerability was each governor's individual approval rating. But given the strong pro-Democratic climate at the time, Democratic incumbents tended not to attract highly experienced Republican challengers regardless of their personal popularity level. As such, there is considerably less variance in challenger experience among Republican challengers than among Democratic challengers (see Brown 2008). This difference alone may explain why experience seems to benefit only Democratic challengers and not Republican ones—there might not be sufficient variation in Republican challengers' experience to estimate the effect of challenger experience against Democratic incumbents. Regardless of whether this partisan difference is real or artifactual, though, these analyses provide no evidence that challenger experience affected election results against Democratic incumbents, though they provide weak evidence that experience mattered against Republican incumbents.

Does Money (alone) Help?

Laying aside any effects of challenger experience on the election result, should we expect challenger spending to matter? In theory, the effects of spending need not be the same on election results as on approval ratings. In gubernatorial elections, potential challengers are not the only strategic actors; potential donors also act strategically. After all, campaign donors prefer not to waste their money contributing to a hopeless campaign. Before contributing, strategic donors consider two factors: The incumbent's vulnerability and the challenger's quality. Challengers raise more money if the incumbent is unpopular, the challenger is well-regarded, or both (Brown 2008). As such, a challenger's fundraising success indicates (in part) his perceived quality—and perceived quality may be a better measure of challenger strength than the experience variables used in the preceding section.

The analyses below follow the same pattern as in the previous section by inserting spending variables into the baseline model to see whether they improve it. Table 9 presents two models showing the effect of the challenger's spending on the election result. The first uses the challengers' total spending, including spending financed personally by the candidate; the second omits self-finance from the challenger's spending total. Both variants improve on the baseline model given earlier. The estimated effect of gubernatorial popularity remains essentially unchanged, but including the spending variables renders partisanship and presidential approval entirely insignificant. The fit rises from an R^2 of 0.59 in the baseline to 0.69 and 0.77 in the models below; adjusted R^2 rises from 0.54 to 0.61 and 0.72, respectively.

[Table 9 about here]

More important than fit, though, are the coefficients themselves. Challenger and incumbent spending are significant in both models. Moreover, the second model has stronger estimated effects and much better fit than the first model; omitting self-finance improves the model. This pattern suggests that spending may have both a direct and an indirect effect on election results. The direct effect is obvious—it takes money to produce advertisements, hire consultants, and run focus groups, and these uses of money can persuade voters to switch sides. The indirect aspect is subtle—regardless of what challengers do with their money, the fact that they can raise it indicates that strategic contributors have faith in the challenger.[15] The direct effect pertains to how the money is used, regardless of the money's source; the indirect effect pertains to the money's source, regardless of how it is used.

Earlier, I used political experience as a proxy for candidate quality. Money raised is a potentially better proxy, since donors take account of much more than political experience; they also consider charisma, policy positions, determination, and likeability. By omitting self-finance

from the challenger's spending total, we give relatively greater weight to this indirect effect in the second model than in the first—and it is this indirect effect that interests us most here, since we are using contributions to the challenger as an empirical indicator of challenger quality. For this reason, the remainder of this section ignores challenger self-finance when discussing campaign spending unless otherwise noted.[16]

Returning to the results in Table 9, observe that challenger and incumbent spending have a similar relationship to one another as they did in the previous section, when the dependent variable was approval. In contrast to the previous section, the difference between the two spending effects is smaller; based on these coefficients, a challenger would need to raise only 30% more than the incumbent in order for the negative effect of the challenger's spending to outweigh the positive effect of the incumbent's. But despite this somewhat lower hurdle, there are still only two challengers who managed to meet this mark when self-finance is excluded. Perhaps as a result, the typical incumbent's share of the two-party vote ended up being 3.0 percentage points higher than his late 2005 approval rating.

As in the previous section, combining challenger and incumbent spending into a single variable helps to clarify the true effects of challenger spending. When challenger self-finance is excluded, the typical challenger raised only 30.8% of the two-candidate total. The weakest challenger raised only 3.1% of the total; the strongest challenger raised 67% of the total, or twice as much as the incumbent.

Re-estimating the previous regression using this new variable produces roughly the same fit as above, along with essentially the same coefficients on the baseline variables (see Table 10).[17] For every percentage point increase in the challenger's spending as a share of the total, the incumbent's expected vote share falls by 0.19 percentage points. In gubernatorial elections,

unlike Congressional elections, what matters is not the challenger's raw finances, but the challenger's ability to compete with the incumbent.

[Table 10 about here]

The graph in Figure 2 depicts these results visually, showing the effect of challenger spending on the incumbent's vote share with out-party gubernatorial approval held at its average (41.5 percent). The lines show predicted values; the points show actual values, with each letter representing the incumbent's partisanship.[18] Predictably (given the pro-Democratic tide), Republican incumbents underperformed Democratic incumbents on election day. It would have taken a 21 percentage point (1.2 standard deviation) increase in the challenger's share of the spending total to match this partisan difference. But the effect of spending is also meaningful; as challengers spent more relative to incumbents, their electoral fortunes improved noticeably.

[Figure 2 about here]

As it turns out, however, these aggregate analyses mask the same partisan pattern observed when examining incumbent approval. Although the figure above shows that challengers of both parties benefited from their spending, more detailed analysis shows that the effect was slightly clearer among Republican challengers to Democratic incumbents. Table 11 presents the relevant estimates. OLS 1 replicates the baseline equation given earlier, but for Democratic incumbents. OLS 2 adds the challenger's spending as a share of the total, with an estimated effect roughly the same as that found in the aggregate model. The fit also improves considerably, even after adjustment for the number of variables.[19] When looking at Republican challengers, then, money clearly affects the electoral outcome.

[Table 11 about here]

By contrast, the challenger's spending has a less clear effect for Democratic challengers

to Republican incumbents (see Table 12). OLS 1 is the baseline model, with OLS 2 specified the same as in the previous table. In OLS 2 the challenger's spending has the same estimated effect as in the previous table but with a larger standard error, rendering the estimate statistically insignificant. This larger standard error appears to arise as a result of some collinearity between Bush approval and spending in states with Republican incumbents ($r=-0.81$, $p<0.001$); removing Bush approval from the model makes the spending variable statistically significant, but at a risk of introducing omitted variable bias.[20] As such, we cannot reject the possibility that Democratic challenger spending has no effect on Republican incumbent vote shares. At the same time, the estimated coefficient is the same for challengers of either party—only the standard error changes—suggesting that a larger-N, multi-year study might find that spending works equally well for challengers of either party.

[Table 12 about here]

This section leads to three general conclusions about the effects of money in gubernatorial elections. First, it appears that money does help gubernatorial challengers—but only to the extent that their spending rises relative to the incumbent's. Second, spending helps the most when we ignore the challenger's self-contributions—suggesting that spending matters on election day more as an indicator of the challenger's credibility than because of its direct effects. And third, spending has the clearest effect in Republican challenges to Democratic incumbents, precisely the opposite pattern (but far less pronounced) as occurs with challenger experience.

Do Experience and Money (combined) Help?

When using experience and spending to predict the change in the incumbent's approval ratings, the effects were additive, not interactive. Experience and spending did not interact with

one another, nor did they interact with the incumbent's late 2005 approval ratings. The same pattern holds true when predicting the incumbent's vote share. The estimated effects reported above are additive, not interactive; combining them into a single model produces no new insights. Even in a combined model, challenger experience continues to have only a small effect (primarily against Republican incumbents) and challenger spending has a much stronger effect (particularly against Democratic incumbents).

Discussion

This paper began with four implicit hypotheses about whether challenger strength might affect the incumbent's approval ratings and vote share:

- That the challenger's experience would hurt the incumbent, either because experience makes the challenger a better campaigner or because experience makes voters more comfortable entrusting the challenger with the governor's office;
- That the challenger's spending would hurt the incumbent, either because it enables the challenger to hire advisors and purchase advertisements or because it serves as an indicator of the challenger's true quality;
- That these variables might interact, either with one another or with the incumbent's initial vulnerability as measured in late 2005;
- Or that neither experience nor spending would matter at all—to the extent that strong challengers outperform weak ones, they do so as a result of the incumbent's initial vulnerability (the null hypothesis).

We can reject the null hypothesis. While it is true that the strongest challengers do take advantage of the incumbent's initial weakness by choosing strategically whether to run, strong

challengers also contribute to that weakness to a small degree. Challenger spending has a negative, statistically significant effect on the incumbent's vote shares and approval. Likewise, challenger experience has a negative, statistically significant (one-tailed) effect on the incumbent's vote shares (but not approval ratings), but only if the incumbent was a Republican in 2006. These effects are not interactive.

Not only does this analysis show that challenger strength does matter, it also helps us understand how it matters. Because the challenger's experience affects vote shares but not approval, we learn that politically experienced challengers were not necessarily better campaigners in 2006; they outperformed inexperienced challengers on election day only because their political experience made them a realistic alternative to the incumbent. In contrast to experience, though, the challenger's spending affects both approval and vote shares, showing that a challenger's ability to raise funds (regardless of her level of previous political experience) indicates her ability to campaign well. Challenger money affects election results because it helps challengers attack the incumbent; challenger experience affects election results because it gives voters a real choice on election day.[21]

All the same, though, the real-world effects of challenger strength are small. True, the challenger's spending hurts the incumbent, but the incumbent's spending has an even stronger effect in the opposite direction. In almost every case, the incumbent managed to outspend the challenger, so the net effect worked in the incumbent's favor. As a result, concluding that challenger spending matters is correct only academically; in the real world, few challengers actually raised enough money to defeat the incumbent. This finding is unusual and unexpected. In Congressional elections, incumbents gain little from their own spending; challenger spending hurts the incumbent, but defensive spending does little to blunt the attack (Jacobson 2004). This

insight does not apply to gubernatorial elections.

This discrepancy between how gubernatorial and Congressional elections operate highlights the need for increased research in the field of gubernatorial elections. Political scientists have spent many fruitful years analyzing the minutest details of Congressional elections, an effort that will surely continue to produce insightful research. Much of what we have learned from the Congressional literature probably applies to other electoral contexts, but we cannot assume that all of it does. Until we look closely at gubernatorial, state legislative, and other subnational elections, we will not know which theories are universal and which apply only to Congress.

We should not be surprised that gubernatorial elections might differ from Congressional ones, given the structural differences between the two types of office. As chief executives of their states, governors are highly visible. They take immediate blame for every bad thing that happens in the state, just as the president's approval suffers for bad things that happen to the nation. When campaign season rolls around, gubernatorial challengers might have trouble telling voters anything bad about the incumbent that voters do not already know. The advantage instead goes to the incumbent, who spends his time telling voters about all the good things he accomplished while they were focused on the short-term problems; the incumbent can also attack the challenger's experience and qualifications. Along these lines, it is telling that most of the rise in approval during the 2006 campaign occurred among the governor's partisans—those most likely to be receptive to new positive information.

On the other hand, members of Congress manage to avoid blame for much of what Congress does from day to day. Because legislating is a collective enterprise, individual members can cast blame for unpopular votes on the rest of the Congress. Meanwhile, they visit

their districts, cut ribbons at new museums, and find other content-free ways of promoting themselves (Mayhew 1974). But when campaign season comes around, shrewd challengers advertise to voters every poorly considered vote that the incumbent has cast. These negative messages provide new information about the incumbent, which the incumbent is hard-pressed to deflect. Along these lines, it is telling that most of the fall in approval during the 2006 Senate campaigns occurred among out-party respondents—those most likely to be receptive to new negative information (Brown and Jacobson 2008).

In short, campaigns give gubernatorial incumbents opportunities for redemption while putting Congressional incumbents at risk of condemnation.[22] While this argument is admittedly speculation, it may help explain why gubernatorial incumbents seem to gain so much from their own spending even though Congressional incumbents gain so little. But regardless of whether these suggestions are accurate, one thing remains certain: We need more research dealing specifically with gubernatorial elections.

---

[1] Experience is the highest percentage of the state the challenger had previously represented in elected office, logged; spending is the challenger's percentage share of the two-candidate spending total. Results available upon request.

[2] Because of the extremely small sample size, I dummy it into two categories rather than preserving all four. Under this specification, 8 of 13 Democratic challengers were experienced, along with 6 of 13 Republican challengers.

[3] To clarify, the first period is the average of monthly surveys from May through December 2005; the second period is the average of two monthly surveys, one in mid-September and one in mid-October. Roughly the same results obtain when using other reasonable sets of months, e.g. when using October alone as the second period.

[4] The in-state change in Bush's approval ratings mentioned earlier does not improve any of the models in this section, so it is omitted entirely.

[5] The estimated coefficient varies in significance and magnitude depending on how experience is measured and on which type of approval data is used (aggregate, out-party, independents, in-party), but under no specification is it significantly negative. This unexpected finding persists when switching from the first-differenced specification to a lag, with October approval on the left and average 2005 approval on the right. There are no outliers responsible for these results.

[6] The correlation between the experience dummy and May-December 2005 approval is -0.55 (p=0.004).

[7] Logged challenger spending ranges from 12.7 to 17.6; logged incumbent spending ranges from 13.4 to 17.6. The standard deviations are 1.4 and 1.1, respectively. Similar results obtain when using the raw spending totals, but with less consistency.

[8] More formally, ln(incumbent spending) = 0.58 * ln(challenger spending) + 6.94.

[9] The two challengers are Wisconsin's Mark Green and Michigan's Dick DeVos, who financed most of his own campaign. A third challenger, Oregon's Ron Saxton, raised 64% more than the incumbent.
[10] That is, (share) = 100 * (challenger) / (challenger + incumbent). Challenger self-contributions are included in this total. Although this formulation requires that challenger spending and incumbent spending have equal but opposite marginal effects, the coefficients in Table 3 are just similar enough to satisfy (very loosely) this requirement.
[11] To be clear, I am subtracting 5.0 – 1.0 for Republican governors, 8.4 – 2.3 for Democrats.
[12] For more on Bayesian model selection, see Raftery (1995).
[13] Using in-party data pushes R^2 down to 0.14. Using approval among independents leads to results between these extremes, but closer to the out-party results than to the in-party results.
[14] It is not surprising that aggregate and out-party data produce similar results, given their high correlation ($r=0.94$, $p<0.0001$); see Brown (2008) for a full discussion of this point.
[15] Or, as alluded to earlier, it indicates that contributors believe that the incumbent is beatable—a possibility controlled for by my other right-hand variables.
[16] This argument provides a theoretical interpretation for some related findings in the Congressional context. There, previous research on challenger self-financing has produced findings consistent with this argument. Although Jon Corzine successfully won a New Jersey senate seat in 2000 after spending a record amount of his own money ($60.2 million), this is atypical; the overwhelming majority of self-financed challenges end in failure (Steen 2006). In fact, self-finance had a negative relationship with votes in 1996-2002 House elections (Alexander 2005).
[17] Bush's out-party approval is dropped from this equation because it contributes nothing to the model. The omission causes an adjustment in the partisan dummy but affects nothing else. Bush's approval is entirely uncorrelated with the variable of interest, challenger spending as a share of the total ($r = -0.06$, $p=0.76$).
[18] Note that the distance between each point and the line is not equal to the residual. The line is drawn with gubernatorial approval held at its average, but approval varies and affects the points.
[19] In addition to R^2 and adjusted R^2, BIC testing also supports OLS 2 over OLS 1.
[20] When Bush's approval is omitted, the estimated effect of spending increases to -0.23 ($p=0.02$). However, the strong negative correlation between Bush approval and challenger spending means that omitting Bush's approval from the model will tend to bias the model toward overstating the effect of spending. (The correlation between Bush approval and spending is far weaker in the aggregate model and in the Democratic incumbent model, so it does not affect those.)
[21] As an extension of this argument, omitting challenger self-finance improves our predictions of election results but not approval ratings because money raised (as opposed to self-financed) serves as an analytical indicator of whether voters were faced with a reasonable alternative on election day.
[22] For a more detailed version of this argument, see Brown and Jacobson (2008).

References

Bond, Jon R., Cary Covington, and Richard Fleisher. 1985. Explaining challenger quality in Congressional elections. *Journal of Politics* 47 (June): 510-529.

Bond, Jon R., Richard Fleisher, and Jeffrey C. Talbert. 1997. Partisan differences in candidate quality in open seat House races, 1976-1994. *Political Research Quarterly* 50 (June): 281-299.

Brown, Adam R. 2007. Gubernatorial approval and strategic entry in the 2006 elections. Presented at the Annual Meeting of the Midwest Political Science Association, Chicago, Illinois, April 2-15.

Brown, Adam R. 2008. *Strategic Politicians in Gubernatorial Elections*. PhD diss., University of California, San Diego.

Brown, Adam R., and Gary C. Jacobson. 2008. Party, performance, and strategic politicians: The dynamics of elections for senator and governor in 2006. Forthcoming in winter issue of *State Politics and Policy Quarterly*.

Goodliffe, Jay. 2001. The effect of war chests on challenger entry in U.S. House elections. *American Journal of Political Science* 45 (October): 830-844.

Goodliffe, Jay. 2007. Campaign war chests and challenger quality in Senate elections. *Legislative Studies Quarterly* 32 (February): 135-156.

Jacobson, Gary C. 1989. Strategic politicians and the dynamics of U.S. House elections, 1946-86. *American Political Science Review* 83 (September): 773-793.

Jacobson, Gary C. 2004. *The politics of Congressional elections*. 6th edition. New York: Longman.

Jacobson, Gary C. 2006. Polarized opinion in the states: Partisan differences in approval ratings of governors, senators, and George W. Bush. Paper prepared for delivery at the Annual Meeting of the Midwest Political Science Association, Chicago, Illinois, April 20-23.

Leal, David L. 2006. *Electing America's governors: The politics of executive elections*. New York: Palgrave Macmillan.

Lublin, David Ian. 1994. Quality, not quantity: Strategic politicians in U.S. Senate elections, 1952-1990. *Journal of Politics* 56 (February): 228-241.

Squire, Peverill. 1992. Challenger profile and gubernatorial elections. Western Political Quarterly 45 (March): 125-142.

Westlye, Mark C. 1991. *Senate elections and campaign intensity*. Baltimore: Johns Hopkins University Press.

Appendix: Tables and Figures

Table 1: Changes in the Incumbent's Approval Rating (Baseline Models)

	All	Republicans	Democrats
Governor is a Republican	-3.46		
	(2.36)		
Constant	4.26*	0.80	4.26*
	(1.67)	(1.75)	(1.58)
N	26	13	13
R^2 (adjusted)	0.08 (0.04)		

Note: Standard errors in parentheses. *p≤0.05.

Table 2: Effects of Challenger Experience on Incumbent's Popularity

	All Incumbents	Republicans	Democrats
Governor is a Republican	-4.27†		
	(2.17)		
Experienced challenger (dummy)	5.32*	6.27†	4.42
	(2.18)	(3.26)	(3.02)
Constant	1.80	-3.06	2.22
	(1.82)	(2.56)	(2.05)
N	26	13	13
R^2 (adjusted)	0.27 (0.21)	0.25 (0.18)	0.16 (0.08)

Note: Standard errors in parentheses. †p≤0.10, *p≤0.05.

Table 3: Effects of Challenger Spending on Incumbent's Popularity

	Including self-finance	Excluding self-finance
Governor is a Republican	-4.32*	-3.59†
	(1.83)	(1.91)
Incumbent's spending (logged)	5.53***	4.70***
	(1.29)	(1.23)
Challenger's spending (logged)	-3.25**	-2.41**
	(0.99)	(0.87)
Constant	-33.13*	-33.56*
	(13.10)	(13.75)
N	26	26
R^2 (adjusted)	0.50 (0.43)	0.45 (0.37)

Note: Standard errors in parentheses. †p≤0.10, *p≤0.05, **p≤0.01, ***p≤0.001.

Table 4: Effects of Challenger Spending on Incumbent's Popularity, by Party

	All Incumbents	Republicans	Democrats
Governor is a Republican	-4.10†		
	(2.05)		
Challenger's share of spending total	-0.18**	-0.14	-0.21**
	(0.06)	(0.11)	(0.06)
Constant	10.80***	5.33	11.83***
	(2.59)	(4.04)	(2.51)
N	26	13	13
R^2 (adjusted)	0.34 (0.29)	0.12 (0.04)	0.51 (0.47)

Note: Standard errors in parentheses. †p≤0.10, *p≤0.05, **p≤0.01, ***p≤0.001.

Table 5: Changes in Gubernatorial Approval, by Party and Spending

	Governor's Partisans			Challenger's Partisans		
	May-Dec 2005	Sep-Oct 2006	Change	May-Dec 2005	Sep-Oct 2006	Change
Republican governors						
All (13 incumbents)	71.8	76.8	5.0	42.4	43.4	1.0
Chal spent < 30% (6)	69.4	74.2	4.8	45.8	47.9	2.1
Chal spent > 30% (7)	73.9	79.1	5.2	39.5	39.6	0.1
Difference in changes			-0.4			2.0
Democratic governors						
All (13 incumbents)	62.6	71.0	8.4	40.5	42.8	2.3
Chal spent < 30% (7)	64.8	75.4	10.7	49.6	55.7	6.2
Chal spent > 30% (6)	60.1	65.9	5.8	30.0	27.8	-2.3
Difference in changes			4.9			8.5

Table 6: The Incumbent's Share of the Two-Party Vote (Baseline Models)

	OLS 1	OLS 2
Governor is a Republican	-30.47*	-45.68†
	(13.51)	(23.52)
Governor's approval (May-Dec 2005)	0.45***	
	(0.09)	
Bush's approval (May-Dec 2005)	0.28†	
	(0.15)	
Governor's out-party approval (May-Dec 2005)		0.33***
		(0.06)
Bush's out-party approval (May-Dec 2005)		0.44†
		(0.25)
Constant	50.19***	84.31***
	(7.16)	(20.08)
N	26	26
R^2 (adjusted)	0.55 (0.49)	0.59 (0.54)

Note: Standard errors in parentheses. Bush's approval is measured at the state level; negative in states with a Democratic governor. †p≤0.10, *p≤0.05, **p≤0.01, ***p≤0.001.

Table 7: Effects of Experience on Republican Incumbents' Vote Share

	OLS 1	OLS 2
Governor's out-party approval (May-Dec 2005)	0.19†	-0.01
	(0.09)	(0.13)
Bush's out-party approval (May-Dec 2005)	0.88*	1.13**
	(0.34)	(0.32)
Experienced challenger (dummy)		-8.10†
		(3.87)
Constant	38.23***	48.24***
	(5.50)	(6.74)
N	13	13
R^2 (adjusted)	0.59 (0.51)	0.73 (0.64)

Note: Standard errors in parentheses. Bush's out-party approval is his state-level approval among the governor's opposition. †p≤0.10, *p≤0.05, **p≤0.01, ***p≤0.001.

Table 8: Effects of Experience on Democratic Incumbents' Vote Share

	OLS 1	OLS 2
Governor's out-party approval (May-Dec 2005)	0.41***	0.40***
	(0.08)	(0.09)
Bush's out-party approval (May-Dec 2005)	-0.12	-0.13
	(0.36)	(0.38)
Experienced challenger (dummy)		-0.41
		(2.41)
Constant	55.33†	56.49†
	(27.07)	(29.29)
N	13	13
R^2 (adjusted)	0.74 (0.69)	0.74 (0.65)

Note: Standard errors in parentheses. The dependent variable is the incumbent's share of the two-party vote. Bush's out-party approval is his state-level approval among the governor's opposition. †p≤0.10, *p≤0.05, **p≤0.01, ***p≤0.001.

Table 9: Effects of Challenger Spending on Incumbent's Vote Share

	Including self-finance	Excluding self-finance
Governor is a Republican	-14.75	-6.81
	(24.71)	(20.79)
Governor's out-party approval (May-Dec 2005)	0.32***	0.29***
	(0.08)	(0.07)
Bush's out-party approval (May-Dec 2005)	0.11	0.03
	(0.26)	(0.22)
Incumbent's spending (logged)	3.56*	3.90**
	(1.44)	(1.11)
Challenger's spending (logged)	-2.47*	-2.99**
	(1.17)	(0.81)
Constant	39.27	35.75
	(30.12)	(25.60)
N	26	26
R^2 (adjusted)	0.69 (0.61)	0.77 (0.72)

Note: Standard errors in parentheses. Bush's out-party approval is his state-level approval among the governor's opposition; negative in states with Democratic governors. *p≤0.05, **p≤0.01, ***p≤0.001.

Table 10: Effects of Challenger Spending on Incumbent's Vote Share, Revisited

	Coefficient	Standard error
Governor is a Republican	-4.03*	1.59
Governor's out-party approval (May-Dec 2005)	0.26***	0.06
Challenger's share of spending total	-0.19***	0.05
Constant	57.58***	3.23
N	26	
R^2 (adjusted)	0.72 (0.69)	

Note: The dependent variable is the incumbent's share of the two-party vote. *p≤0.05, **p≤0.01, ***p≤0.001.

Table 11: Effects of Spending on Democratic Incumbents' Vote Share

	OLS 1	OLS 2
Governor's out-party approval (May-Dec 2005)	0.41***	0.30**
	(0.08)	(0.08)
Bush's out-party approval (May-Dec 2005)	-0.12	0.12
	(0.36)	(0.30)
Challenger's share of spending total		-0.14*
		(0.06)
Constant	55.33†	45.26†
	(27.07)	(22.26)
N	13	13
R^2 (adjusted)	0.74 (0.69)	0.85 (0.79)

Note: Standard errors in parentheses. Bush's out-party approval is his state-level approval among the governor's opposition. †p≤0.10, *p≤0.05, **p≤0.01, ***p≤0.001.

Table 12: Effects of Spending on Republican Incumbents' Vote Share

	OLS 1	OLS 2
Governor's out-party approval (May-Dec 2005)	0.19†	0.21†
	(0.09)	(0.10)
Bush's out-party approval (May-Dec 2005)	0.88*	0.45
	(0.34)	(0.57)
Challenger's share of spending total		-0.14
		(0.15)
Constant	38.23***	47.9**
	(5.50)	(11.94)
N	13	13
R^2 (adjusted)	0.59 (0.51)	0.63 (0.51)

Note: Standard errors in parentheses. Bush's out-party approval is his state-level approval among the governor's opposition. †p≤0.10, *p≤0.05, **p≤0.01, ***p≤0.001.

Figure 1: Regression to the Mean in Approval Data (Republican Incumbents)

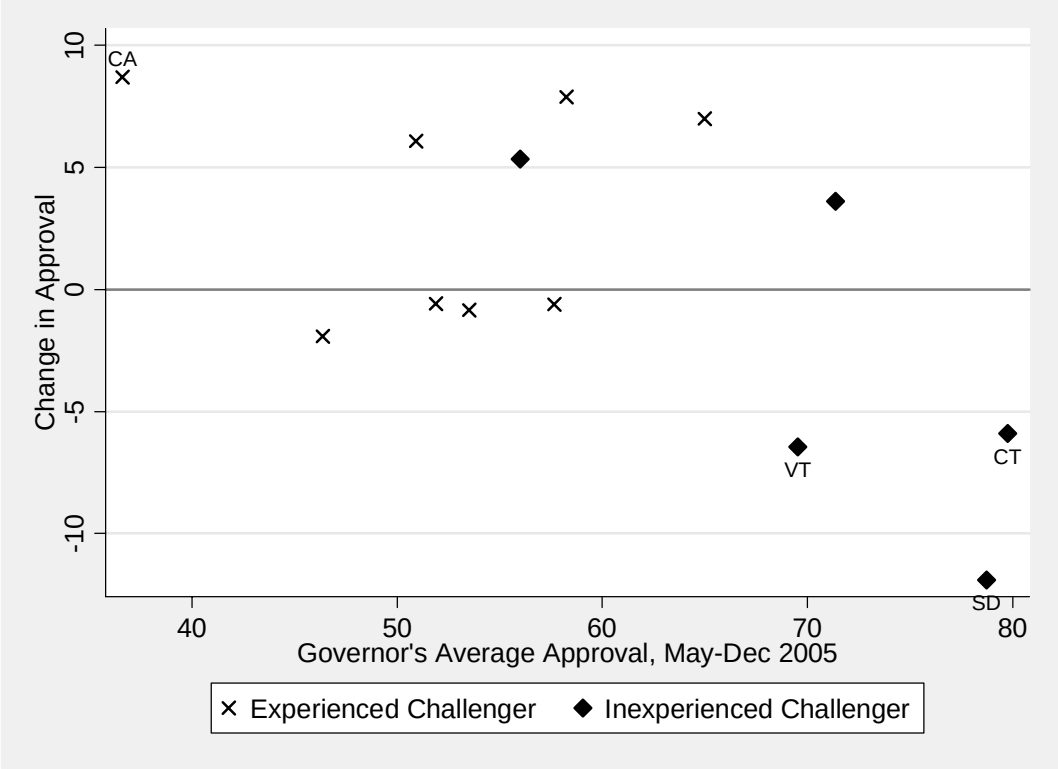

Figure 2: Effects of Challenger Spending on Incumbent's Vote Share