Gubernatorial Approval and Strategic Entry in the 2006 Elections

Adam R. Brown

a10brown@ucsd.edu

Dept. of Political Science
University of California, San Diego
9500 Gilman Drive, M/C 0521
La Jolla, CA 92093-0521

Paper prepared for the 2007 meeting of the Midwest Political Science Association, Chicago, Illinois, April 12-15.

Though it is now widely accepted that candidate quality and strategic donors mediate Congressional election results, this insight has had little treatment in the literature on gubernatorial approval and elections. Rather than examine challengers and donors, most studies have attributed gubernatorial election outcomes entirely to voter behavior, which has the misleading implication that the variables driving election results also determine gubernatorial job approval ratings. This paper presents a new model of gubernatorial elections, paying special attention to how potential challengers in 2006 responded strategically to January gubernatorial approval ratings.

1 Introduction

Politicians are strategic.[1] They do not risk losing a safe seat in the U.S. Senate to challenge a popular governor's reelection bid. Before primary election season rolls around, potential gubernatorial challengers sniff the political winds. They want to know whether it will be a good year for their party and whether it will be a bad year for the incumbent. As a result, we should be able to predict how "good" the governor's challenger will be—and how much money that challenger will raise—by looking at presidential and gubernatorial approval ratings, national and state economic performance, and state-level partisanship and ideology.

Most previous studies of gubernatorial elections have failed to appreciate the importance of this strategic behavior by challengers and donors.[2] Instead, scholars have attributed electoral swings to a desire by voters to "send a message to the incumbent president's party" (Stein 1990, 51). Contrast this with the literature on Congressional elections, where it is now widely accepted that the effects of economic and political conditions on Congressional races are mediated by strategic decisions by politicians, donors, and activists.

In this paper, I explore how state and national factors influence candidate recruitment in gubernatorial elections. I begin with a brief review of what we already know from the literature, after which I present evidence that potential gubernatorial candidates, like House candidates, do behave strategically.

[1] Many thanks to Gary Jacobson for sharing his time, data, and expertise. Equal thanks to Janelle for changing more than her share of diapers so I could work on this project. I also thank Mike Binder, Matt Childers, Craig Burnett, Ricardo Gomez-Vilchis, and Keith Poole for their insights.
[2] Leal (2006) and Squire (1992) are notable exceptions, discussed below.

2 The Literature's Received Wisdom

Although we know a great deal about federal elections, few have examined voting and public opinion at the subnational level. This gap in the literature ought to concern us; after all, most of a voter's choices on Election Day relate to state, county, and local politics. Even gubernatorial elections have received scant treatment, and most of it recent.[3] We know only a few basics. For example, we know that incumbency matters (Ansolabehere and Snyder 2002; Turett 1971; Tompkins 1984), as do each state's partisan leanings (Chubb 1988), although the effect is slight (Erikson et al. 1993).

But we do not know whether voters apply the same criteria when evaluating subnational politicians—governors in particular—that they use when evaluating national ones. Instead, we observe two competing hypotheses. The first hypothesis argues that governors need worry about national partisan tides and little else (Crew and Weiher 1996; Peltzman 1987), particularly if the governor belongs to the president's party (Stein 1990). The second hypothesis allows that national tides can matter, but suggests that voters also use federalism as a voting cue; that is, they judge national politicians based on their foreign policy and macroeconomic records, but they judge state and local politicians based on tax rates, potholes, and school quality (Atkeson and Partin 1995, 2001; Carsey and Wright 1998; Niemi et al. 1995).

Evidence for this second (state-level) hypothesis has been mixed. Although recent scholarship suggests that state economies appear to have grown increasingly independent of national trends (Brace 1993), this finding does not immediately imply that gubernatorial election results will reflect state economic trends. As it turns out, the

[3] Studies of this topic began in earnest with Kenney (1983), Holbrook-Provow (1987), Peltzman (1987), Chubb (1988), and Simon (1989).

literature's findings have depended on how state economic performance is measured. Incumbents lose when they raise taxes or expand the state's budget (Niemi et al. 1995; Peltzman 1987). And when survey respondents are asked to evaluate economic conditions, their evaluations correlate well with incumbent support (Atkeson and Partin 1995; Carsey and Wright 1998; Stein 1990). But raw macroeconomic indicators have inconsistent and possibly contingent effects[4] (see Crew and Weiher 1996; Ebeid and Rodden 2006; Jacobson 2006; Kenney 1983; Leyden and Borrelli 1995; Peltzman 1987).

By contrast, almost every analysis agrees that gubernatorial elections reflect national partisan swings. Governors of the president's party fare better when the president is popular (Carsey and Wright 1998, Crew and Weiher 1996, Niemi et al. 1995), when national unemployment is low (Crew and Weiher 1996, Leyden and Borrelli 1995, Peltzman 1987), and when national economic growth is robust (Chubb 1998, Peltzman 1987). Those few researchers who have challenged these findings have met strong rebuttals (compare Atkeson and Partin 1995, 1998 with Carsey and Wright 1998). Consistent with these findings, there is a strong correlation between postwar gubernatorial, House, Senate, and presidential election outcomes (see Table 2-1). Gubernatorial swings correlate with House swings at 0.74 ($p<0.0001$) and with Senate swings at 0.57 ($p=0.0008$).

[4] Leyden and Borrelli (1995) suggest that macroeconomic indicators matter less under divided government, and Ebeid and Rodden (2006) argue that these indicators matter less in states dependent on primary products (farming, logging, mining).

Table 2-1: Republican Gains and Losses, 1946-2006

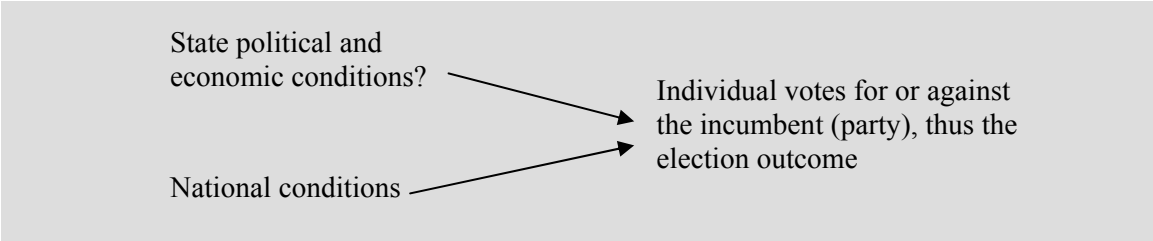
Year	Governors	House	Senate	Presidency
1946	6	56	13	(D)
1948	-8	-75	-9	D
1950	11	28	5	
1952	8	22	1	R
1954	-3	-19	-2	
1956	-2	-2	-1	R
1958	-18	-49	-15	
1960	-5	22	2	D
1962	-7	1	-3	
1964	-10	-37	-1	D
1966	13	47	4	
1968	4	5	6	R
1970	-9	-12	2	
1972	-5	12	-2	R
1974	-20	-49	-4	
1976	-4	-1	0	D
1978	-6	15	3	
1980	1	33	12	R
1982	-16	-26	1	
1984	3	14	-2	R
1986	-2	-5	-8	
1988	2	-2	0	R
1990	-4	-8	-1	
1992	-4	10	0	D
1994	13	54	8	
1996	-3	-3	2	D
1998	12	-4	0	
2000	-5	-1	-5	R
2002	8	8	1	
2004	0	3	4	R
2006	-6	-29	-6	

Note: I have listed the net change for Republicans relative to the Democrats only; negative numbers indicate a net Republican loss, and third-party victors are excluded. *Presidency* gives the victor's party in presidential election years. Some gubernatorial elections are held in odd-numbered years; these elections are excluded.
Sources: Gubernatorial data through 1998 from Rusk (2001, Table 7-2). House and Senate data through 2002 from Jacobson (2004, Table 6-1). Data for later years compiled by author.

Still, how should we interpret the literature's findings about gubernatorial elections? In post-analysis interpretations, we often read that citizens punish or reward

their governors in order to send a message to Washington, "to 'settle up' with the president for the past two years" (Peltzman 1987, 296): "For many citizens, political judgments are general indictments or rewards of the party in [national] power" (Carsey and Wright 1998). Even the studies that argue for state-level variables speak as though nothing intervenes between voters and their incumbent. Figure 2-1 presents this implicit model graphically.

Figure 2-1: The Literature's Model of Gubernatorial Elections



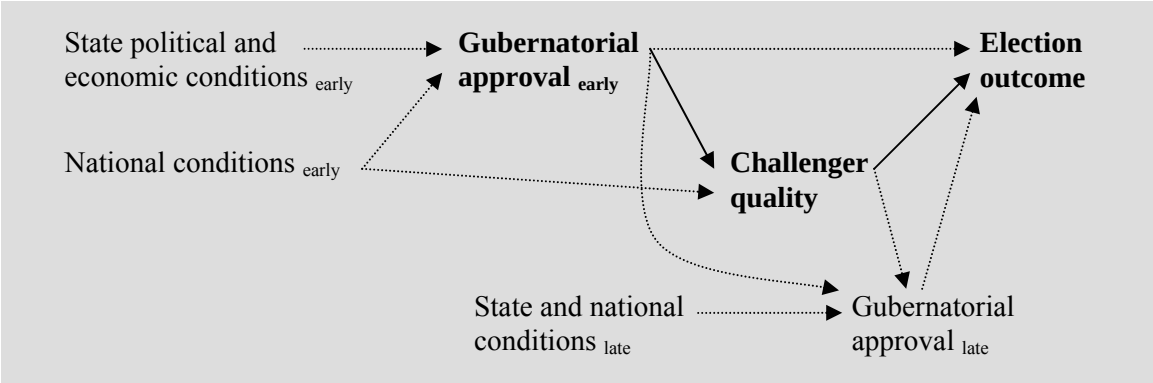
The literature on Congressional elections used a similar model thirty years ago when interpreting accumulating evidence that national conditions strongly affect Congressional elections (Kramer 1971; Tufte 1978). As Tufte wrote, "The vote cast in midterm congressional elections is a referendum on the performance of the President and his administration's management of the economy" (1975, 824). Trouble arose, though, when survey researchers discovered that individual voters did not generally link their economic situation with their Congressional votes (Fiorina 1978; Hibbing and Alford 1981), except under highly constrained circumstances (Weatherford 1978).

Eventually, Jacobson and Kernell (1983) reminded us that elections are about more than just incumbents and voters—challengers and donors act strategically and play a critical role by influencing the choices available to voters on Election Day. Politicians and those who finance and recruit them read political and economic conditions to estimate their party's chances of success in the coming election. In Congressional

elections, "The strategic decisions of politicians so structure the vote choice that electoral results are consonant with national level forces even if individual voting decisions are not" (Jacobson and Kernell 1983, 3).

Applying these arguments to the gubernatorial context leads to the updated model of gubernatorial elections shown in Figure 2-2—a model in which economic and political factors early in the election year affect the incumbent's popularity, which determines the quality of the governor's eventual challenger, which—together with economic and political factors later in the year—influences the ultimate outcome of the election.[5] The steps of greatest theoretical interest to this paper are indicated with bold type and solid arrows.

Figure 2-2: An Updated Model of Gubernatorial Elections



In two papers currently under development, I explore the first and last steps in this causal chain—that is, I study the variables driving gubernatorial approval and then I ask whether challenger quality actually affects election outcomes. Here, though, I focus on candidate recruitment as a dependent variable, demonstrating a simple model that can explain challenger quality, at least in the 2006 gubernatorial elections.

[5] To keep the model readable, Figure 2-2 omits arrows connecting early state and national conditions to late conditions.

3 Theory

The existing literature has paid scant attention to challenger quality.[6] The most comprehensive treatment comes from Leal (2006), whose otherwise thoughtful analysis suffers from a crucial shortcoming that I return to shortly. First, though, consider how strategic challengers might be expected to behave in gubernatorial elections. We often model voting behavior using a rational calculus (Riker and Ordeshook 1968). Because candidates, activists, and donors also behave rationally, we can model their behavior using an analogous calculus. Following Jacobson and Kernell's lead, I apply Black's (1972) calculus of the upwardly mobile politician to the gubernatorial case. Per Black,

$$U_0 = (PB) - R,$$

where U represents the utility to a politician of seeking office *0*, P indicates the probability of winning an election to this office, B signifies the benefits of holding the office, and R denotes the risks inherent in mounting a campaign.[7] If U_0 is positive, the politician runs for office *0*.

Risk is fundamental to this calculus, for political office represents not only a prize but a resource. Officeholders acquire contacts, recognition, experience, and tenure, all of which strengthen their ability to mount future political campaigns. But cashing in on these resources also places them at risk, since an unsuccessful campaign can force a politician out of politics. As Leal writes, "It appears that after running an unsuccessful statewide campaign, candidates either leave politics or run for lower office" (2006, 25).

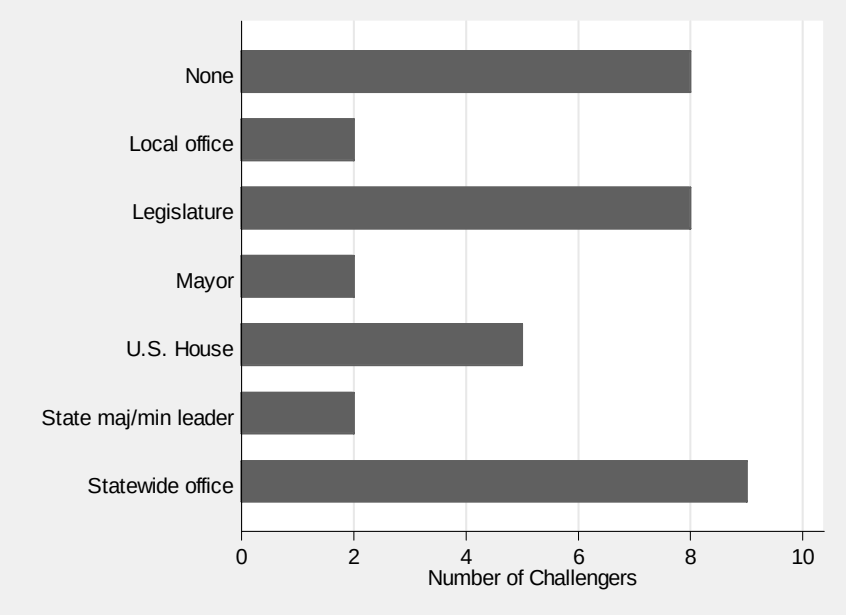
---

[6] Some earlier analyses did treat this topic, but incompletely. Niemi et al. (1995), for example, consider whether incumbents are more likely to seek reelection when times are good, but they do not consider whether high-quality challengers are more likely to enter the fray when times are bad. And Squire (1992) made an explicit attempt to replicate Jacobson and Kernell's analysis at the gubernatorial level, but his results are undermined by his use of an unpersuasive index of challenger quality.

[7] Black's formulation used *C*, representing costs, instead of *R*; the result is the same.

This logic applies even to retired officeholders considering a reentry into politics, since an unsuccessful run can disenchant donors and activists. Since R increases with higher offices, U_0 will be negative for experienced politicians unless PB is high. And as we would expect, candidate quality does vary considerably. In 2006, for instance, there was everything from a retired governor to a few obscure attorneys seeking election (see Figure 3-1).

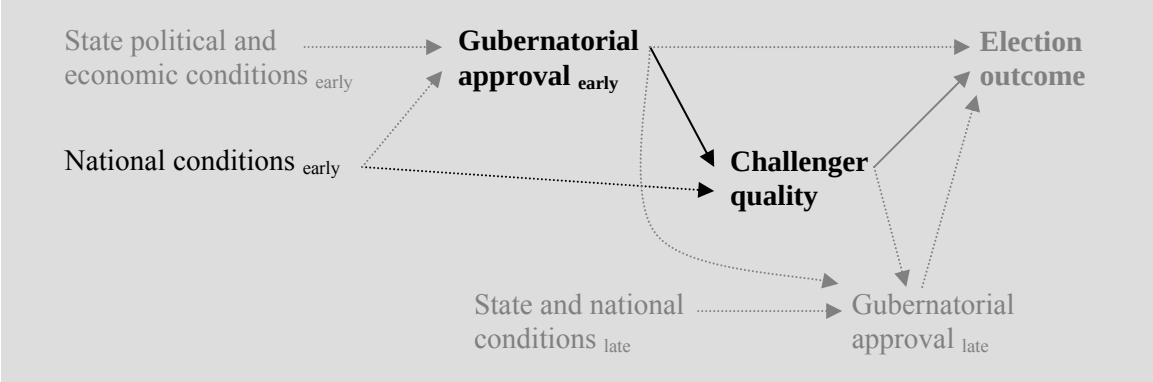
Figure 3-1: Challenger Experience in the 2006 Gubernatorial Elections



Wary of these risks, potential challengers estimate their probability of success by looking for signs that the incumbent will be weak in November. Broadly speaking, challengers consider the governor's personal popularity and the national partisan mood to estimate P. With few exceptions, the variables used in previous analyses of gubernatorial elections could be construed as indicators of these two conceptual variables. By using direct measures of these two concepts, however, I eliminate the need for including many of the indicators used in the literature. I measure gubernatorial approval in January and use a simple partisan dummy to record the national partisan mood in 2006.

Some might question whether controlling for the governor's partisanship is necessary; after all, to the extent that national conditions matter, they are probably absorbed (imperfectly) into gubernatorial approval. Nevertheless, national conditions could have an independent effect if potential challengers use them to estimate the direction in which gubernatorial approval might move in the future, so they remain in the model. This discussion leaves us with the reduced model shown with dark print in Figure 3-2.

Figure 3-2: Modelling Challenger Quality



Challengers weight their probability of victory against the benefits of winning, B. In House elections, B washes out since all House seats are of equal value. B probably washes out in gubernatorial elections as well. While we might expect B to be higher in larger or wealthier states—since governors of these states can redistribute vaster resources—there are also more mayors, House members, and other officials that might seek the governorship, offsetting any increase in B with a decrease in P. These considerations are testable, though I leave them to future analysis on account of the present study's small sample size.

4 Data and Methods

Operationalizing candidate quality is always difficult. Candidates who have represented the entire state are clearly high-quality challengers. This includes former governors as well as state legislative majority (or minority) leaders. In 2006, eleven challengers had statewide experience. And candidates who have held no electoral experience are almost always low-quality challengers, unless they have some other source of name recognition. In 2006, for example, Pennsylvania's challenger had wide name recognition from his days with the Pittsburgh Steelers, and Massachusetts's challenger had previously gained prominence as a Clinton appointee. Excluding these two, six inexperienced challengers remain as low-quality candidates.

But between these two extremes are a variety of politicians who have held some level of electoral office, as shown in Table 4-1. At one end, we have Iowa's and Oklahoma's challengers, members of Congress whose districts covered twenty percent of their respective states. At the other end, we have New Hampshire's challenger, whose state legislative district covered only 0.25 percent of the state. But it is not obvious how much of the state a challenger must have represented for her electoral experience to be meaningful.

Table 4-1: Coding Challenger Quality

State	Challenger's Highest Electoral Office	Percent of State	Quality
Alabama	Lt Governor	100	High
Alaska	Governor	100	High
Arizona	None (President, Center for Ariz Policy)	0	Low
Arkansas	State attorney general	100	High
California	State treasurer	100	High
Colorado	District Attorney	13.4	Mid
Connecticut	New Haven Mayor	3.7	Low
Florida	State house majority leader	100	High
Georgia	Lt Governor	100	High
Hawaii	State senate whip	4	Mid
Idaho	None (newspaper publisher, attorney)	0	Low
Illinois	State treasurer	100	High
Iowa	U.S. House	20	Mid
Kansas	State senate	2.5	Low
Maine	Assistant state senate leader	2.9	Mid
Maryland	Baltimore Mayor	11.5	Mid
Massachusetts	None (Asst. US Attorney General)	0	Mid
Michigan	None (corporate president)	0	Low
Minnesota	State attorney general	100	High
Nebraska	None (attorney)	0	Low
Nevada	State senate minority leader	100	High
New Hampshire	State representative	0.3	Low
New Mexico	None (state party chair)[8]	0	Low
New York	State attorney general	100	High
Ohio	U.S. House[9]	5.6	Mid
Oklahoma	U.S. House	20	Mid
Oregon	Portland school board	2.3	Low
Pennsylvania	None (Pittsburgh Steelers Hall of Fame)	0	Mid
Rhode Island	Lt Governor	100	High
South Carolina	State senate	2.2	Low
South Dakota	State representative	1	Low
Tennessee	State senate whip	3	Mid
Texas	Houston city council (also U.S. House)[10]	9.7	Mid
Vermont	State senate	3.3	Low
Wisconsin	U.S. House	12.5	Mid
Wyoming	None (attorney, rancher)	0	Low

Coding of "percent of state": For legislative districts, this is the reciprocal of the number of districts. For city-wide offices, this is the percent of the state population residing within city limits. For city council and board districts, I divide the city-wide measure by the number of districts. Rounded to the nearest tenth.

[8] Appointed after the original party nominee dropped out for lack of funds.
[9] As a result of redistricting, Ted Strickland had actually represented much more than 5.6% over the years.
[10] A Houston city council district includes more of the population than a Texas House district.

In 2006, seventeen challengers fit into this middle category. Of them, ten had represented less than five percent of the state; two had represented between five and ten percent of the state; three had represented between ten and fifteen percent of the state; and two had represented twenty percent of the state. For ease of analysis, I divide these seventeen politicians into two groups. The ten who had represented less then five percent of the state I classify as low quality candidates; the rest, I classify as medium quality candidates. I also code the challengers from Pennsylvania and Massachusetts as medium quality for the reasons discussed above. This cutoff (coincidentally) places all the House members in the mid-quality category and all the state legislators into the lower category, which fits our intuition that these offices provide fewer resources for a gubernatorial run. As part of my analysis, I vary this cutoff to confirm that my findings are robust to its specific placement.

State legislative leadership complicates this classification, though. Clearly, majority and minority leaders are high-quality challengers whose influence permeates the entire state; I list them as representing 100 percent of the state. But state legislative whips and assistant floor leaders lack the same prominence. The small size of their legislative districts suggests that they ought to be considered low-quality candidates, but their leadership roles suggest that they have more going for them politically than other legislators might. To account for these competing considerations, I code these second-tier legislative leaders as mid-quality. With this classification scheme, there are thirteen low-quality candidates, twelve mid-quality candidates, and eleven high-quality candidates.

Candidate quality should reflect gubernatorial popularity and national partisan tides as measured in early 2006. Only recently have good gubernatorial approval data

become available. Previous studies have instead used a variety of proxy measures, though these measures are not generally presented as proxies for gubernatorial approval. For example, Leal's (2006) study regressed challenger quality on the incumbent's age, primary election margin, length of tenure, and previous election margin, as well as state income growth, crime rates, change in SAT scores, and unemployment rates—and found none of them to be significant. Had Leal enjoyed access to a direct measure of gubernatorial approval, these indicators would not have been necessary—and without all these overlapping indicators inflating his errors, Leal might have had different results.

I use only a direct measure of gubernatorial approval (disapproval, actually) and the partisan dummy already discussed. Because I examine only 2006, the partisan dummy captures any national partisan tides. The gubernatorial disapproval variable subsumes other economic and political factors, since these factors matter only to the extent that they hurt the incumbent's popularity.[11] Since dissatisfied voters turn out in larger numbers than satisfied voters (Kernell 1977), would-be challengers ought to care more about disapproval than approval, although these two ratings correlate so closely that essentially the same pattern emerges either way. If this simple analysis can demonstrate that challenger quality reflects the incumbent's approval ratings, then work on gubernatorial approval will become all the more interesting.

5 Findings

In November 2006, thirty-six states held gubernatorial elections. Table 5-1 lists each challenger's quality and each incumbent's January 2006 disapproval ratings for

[11] Gubernatorial approval data is measured monthly by SurveyUSA; see appendix for details.

these states, omitting five states where term limits created open seats and another two where governors left office to pursue more ambitious plans.[12]

Table 5-1: Challenger Quality and Incumbent's January 2006 Approval

	Incumbent Running?	Challenger Quality	Disapproval
Alaska	Lost primary	High	68
California	Yes	High	62
Maine	Yes	Mid	56
Illinois	Yes	High	53
Michigan	Yes	Low	53
Oregon	Yes	Low	48
New York	Retiring	High	47
Wisconsin	Yes	Mid	45
Minnesota	Yes	High	44
Maryland	Yes	Mid	43
Pennsylvania	Yes	Mid	43
South Carolina	Yes	Low	42
Texas	Yes	Mid	41
Alabama	Yes	High	40
Iowa	Retiring	Mid	39
Tennessee	Yes	Mid	38
Rhode Island	Yes	High	36
Georgia	Yes	High	34
New Mexico	Yes	Low	34
Arizona	Yes	Low	32
Hawaii	Yes	Mid	32
Kansas	Yes	Low	31
Oklahoma	Yes	Mid	27
Wyoming	Yes	Low	27
Nebraska	Yes	Low	26
New Hampshire	Yes	Low	25
Vermont	Yes	Low	25
South Dakota	Yes	Low	21
Connecticut	Yes	Low	17

Source: Approval ratings gathered by SurveyUSA. Other data compiled by author.

[12] Mitt Romney left Massachusetts to pursue the presidency, and Idaho's governor left shortly after primary season to fill a federal appointment. I do not omit New York and Iowa, the two other states with retiring governors, since an incumbent's retirement can be a strategic response to the same factors that motivate challengers.

With the states in this table sorted by the incumbent's January disapproval ratings, it is apparent that popular incumbents rarely face experienced challengers. Seventy-five percent of low-quality candidates challenged a governor with disapproval below 34 percent, while seventy-five percent of high-quality candidates challenged a governor with disapproval above 40 percent. And the three states that look like they might be outliers above—Maine, Michigan, and Oregon—have Democratic governors, reflecting our expectation that Democrats would not attract high-quality challengers in 2006.

Table 5-2 presents a simple ordered logit model of challenger quality using this data. Both gubernatorial disapproval and partisanship have significant effects in the expected directions, as suggested by the previous table.

Table 5-2: A Simple Ordinal Model of Challenger Quality

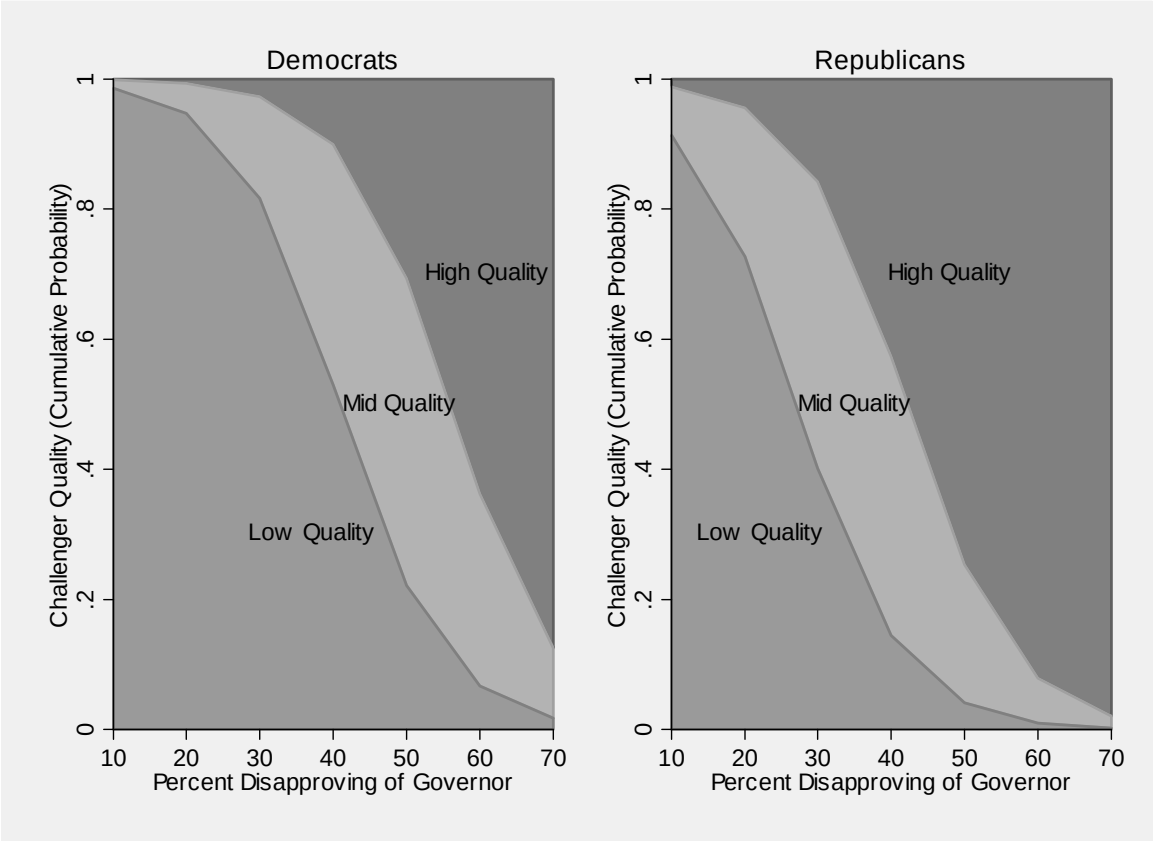
	Coefficient	S.E.
Governor's Disapproval	0.138**	0.046
Governor is a Republican	1.893*	0.875
Cut 1	5.629	1.920
Cut 2	7.703	2.196
N	29	
Log likelihood	-23.206	
Probability	0.0003	
Percent Modal	41.4%	
Percent Correctly Classified	55.2%	
Proportional Reduction in Error	0.235	

Note: The dependent variable is a trichotomous measure of challenger quality, as defined in Table 5-1. See text for variable definitions. Calculation of proportional reduction in error: (percent correct – percent modal) / (1 – percent modal). *p≤0.05, **p≤0.01.

Using these regression results, Figure 5-1 plots the cumulative probability that a governor will face a higher quality challenger against the observed range of disapproval

ratings.[13] In both panels, movement in the incumbent's disapproval ratings rapidly changes the predicted probability that the incumbent will face a high-quality challenger. And across the board, Republicans operate at a clear disadvantage.

Figure 5-1: Predicted Probabilities of Challenger Quality



These analyses use disapproval data from January 2006. With a primary election schedule stretching from March 7 to September 23 (see Table 5-3), it is difficult to know which month's data to use. Perhaps data from May ought to be used, or from three months prior to each state's primary. The results above are robust to either change.

[13] Calculated using the SPOST package for Stata from Long and Freese (2006).

Table 5-3: Primary Election Schedule in 2006

Month	Mar	Apr	May	Jun	Jul	Aug	Sep
Primaries	2	0	6	7	2	8	11

Source: Federal Voter Assistance Program (http://www.fvap.gov/pubs/electcal.html).

These results are also robust to a variety of other manipulations, such as including additional variables,[14] changing the cutoffs for challenger quality, collapsing the dependent variable into various binary measures, and substituting approval for disapproval. Partisanship is significant in only some of these alternative specifications, but the consistent statistical significance of gubernatorial disapproval in all these models supports my claim that other variables enter the strategic calculus only through their effects on incumbent popularity. Although Leal found a variety of state-level variables not to significantly affect challenger quality, it appears that potential challengers do act strategically. The robustness of this result to alternative specifications should partially assuage concerns about this study's small sample size.

6 Concluding Remarks

The existing literature on gubernatorial elections has uncovered many interesting correlations between economic and political variables and election results. But most researchers have overlooked the importance of challengers and donors, whose strategic decisions mediate election outcomes by structuring the choices available to voters on Election Day. This paper has demonstrated that potential gubernatorial challengers do

[14] State-level presidential approval (interacted with partisanship) and state partisan propensities (a measure of how frequently each state has elected politicians from the challenger's party in recent gubernatorial, senatorial, and presidential elections) were tested under several specifications. They were rarely significant. Bayesian model selection supports omitting them (Raftery 1995).

appear to respond strategically to incumbent popularity. Naturally, such a conclusion begs the questions of what drives gubernatorial popularity and whether candidate quality even matters, questions I address in other papers currently under development. But even without knowing precisely which political and economic factors are feeding into gubernatorial approval and thereby influencing candidate recruitment, this paper supports the claim that potential challengers do respond strategically to the governor's popularity.

7 Appendix: Can We Trust SurveyUSA?

Critics have voiced skepticism about SurveyUSA's "Interactive Voice Response" (IVR) methodology, which produces the gubernatorial approval data used in this paper. Under this method, questions are recorded by professional announcers—typically a local television anchor—and respondents answer questions by pressing telephone keys. In principle, anyone can answer the phone and hit the buttons; for this reason, the *New York Times* prohibits its writers from reporting IVR results because "results of this type of poll are not reliable," though this claim has only theoretical justification.[15] Critics have claimed that IVR is useful only to journalists desiring to publish "throwaway factoids of dubious value with the suggestion that they reflect real public opinion" (Traugott 1995).

An empirical analysis of SurveyUSA's record suggests that these theoretical concerns may be overblown. While it is true that SurveyUSA's response rates and cooperation rates are lower than the industry average, they fall within the industry range (Holbrook, Krosnick, and Pfent 2003). More to the point, SurveyUSA's election projections perform well, its job approval ratings track closely to ratings from other pollsters, and its month-to-month data have strong enough internal consistency to discount any theoretical concerns about the IVR methodology.

7.1.1 Election Projections

Most survey data can never be validated, since we do not usually know the "true" percentage of the population espousing a particular view. But gubernatorial and presidential performance data are an exception; unlike most survey data, projections can

[15] From the *New York Times* polling standards, published in June 2006, an internal document describing editorial policies. Available at http://www.nytimes.com/packages/pdf/politics/pollingstandards.pdf.

be verified every few years at election time. During the final months before each election, SurveyUSA asks respondents who they intend to vote for. SurveyUSA is not alone in asking these questions, making a direct comparison with other pollsters possible.

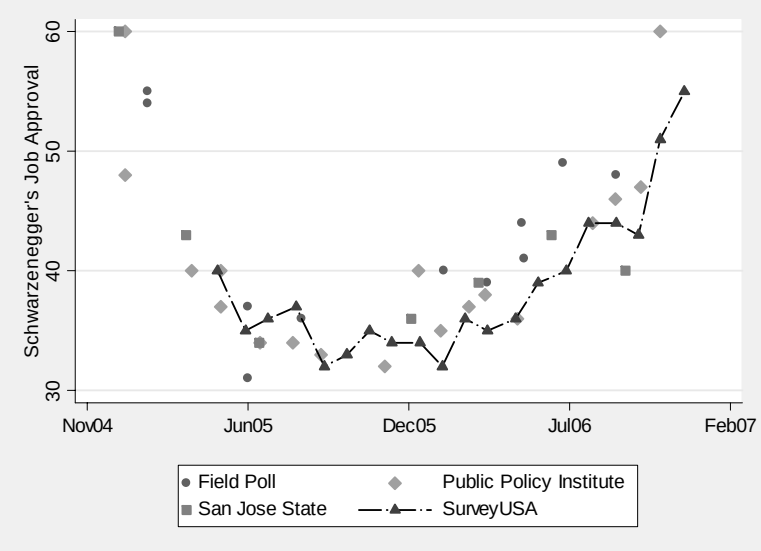
Based on such a comparison, Bloom reports that SurveyUSA performed slightly better than other nonpartisan polling organizations at projecting the 2002 election results, a feat repeated in 2004 (Bloom 2003; Bloom and Pearson 2005). For these tests, Bloom uses a simple but telling measure: Whether 95 percent of election outcomes actually fell within SurveyUSA's projected margin of error. While SurveyUSA fell just short of this standard, it came much closer to it than Gallup and other established firms did.

A separate analysis by Kenner and Saletan comes to similar conclusions. These authors apply the "sum" and "spread" criteria to projections of the 2004 elections. Under the "sum" method, Kenner and Saletan add the error for each candidate in each poll. For example, if Bush won 48% in a state and Kerry won 54%, but SurveyUSA predicted 47% and 52%, the sum of these errors was (48-47) + (54-52), or 3. Under the "spread" method, Kenner and Saletan measure the difference between SurveyUSA's projected spread and the actual spread. In the preceding example, SurveyUSA would have underpredicted the spread between Kerry and Bush by 1 point. In 2004, SurveyUSA and Rasmussen, both of which use IVR, beat Gallup by both the sum and spread methods, and SurveyUSA also performed well in comparison to Zogby and Mason-Dixon (Kenner and Saletan 2004). In Bloom's words, "As much as academic survey researchers may have wished to see SurveyUSA under-perform the field, they clearly did not, and may have actually done better than average" (2003, 15).

7.1.2 Job Approval Data

SurveyUSA's job approval data are more difficult to validate, for the simple reason that few pollsters measure gubernatorial approval. SurveyUSA collected monthly approval ratings in every state from May 2005 until November 2006. During this period, other pollsters collected enough data for a comparison in only three states, according to data compiled by Niemi, Beyle, and Sigelman.[16] In California, SurveyUSA's data competes with the respected Field Poll, the Public Policy Institute of California, and San Jose State University (see Figure 7-1). All four pollsters identify similar trends.

Figure 7-1: SurveyUSA's Record in California

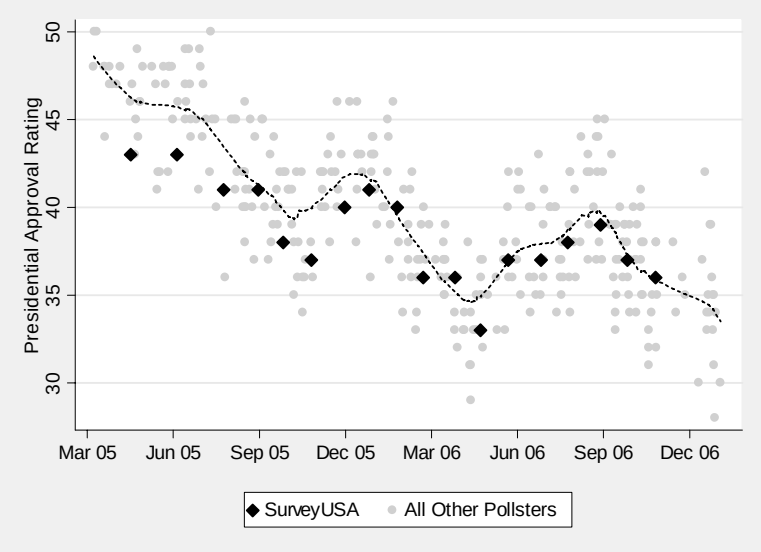


Comparisons are also possible in New York and New Jersey. In both cases, SurveyUSA's data follow other approval ratings closely.[17] SurveyUSA also gathers presidential approval ratings. But unlike most pollsters, SurveyUSA calculates presidential approval ratings by measuring approval independently in every state, then weighting the findings to produce a national result. Figure 7-2 plots SurveyUSA's ratings

[16] With National Science Foundation support, these researchers have compiled decades' worth of approval data. I use the January 15, 2007, version of their data. See http://www.unc.edu/~beyle/jars.html.
[17] Figures available upon request.

against those measured by all other pollsters. The dotted line shows the Lowess-smoothed trend of all other pollsters' data.[18] SurveyUSA's approval ratings fall close to this curve even though they were not used in its calculation. Neither of these figures gives any reason to question whether SurveyUSA's data are valid.

Figure 7-2: SurveyUSA's Record with Presidential Approval



7.1.3 Internal Consistency

SurveyUSA's data also appear internally consistent. Each state's monthly gubernatorial approval ratings correlate well with the preceding month's data, as we would expect. And pairwise correlations with data measured in November 2006 become increasingly weak as we look further back in time, also as we would expect (Table 7-1).

[18] Presidential approval data for other pollsters downloaded from the Roper Center's archives (http://137.99.36.203/CFIDE/roper/presidential/webroot/presidential_rating.cfm) on January 29, 2007. Lowess smoothing uses a bandwidth of 0.15.

Table 7-1: Internal Consistency of Gubernatorial Approval Data

Month	Correlation with Preceding Month	Correlation with November 2006
Nov 06	0.96	1.00
Oct 06	0.97	0.96
Sep 06	0.96	0.95
Aug 06	0.97	0.93
Jul 06	0.96	0.94
Jun 06	0.97	0.93
May 06	0.96	0.92
Apr 06	0.95	0.89
Mar 06	0.95	0.90
Feb 06	0.95	0.89
Jan 06	0.96	0.85

Moreover, Table 7-2 shows that compositional variables remain stable over time. The percentage of respondents in each state calling themselves "pro-life" or "pro-choice" changes little from month to month. Partisanship is slightly more variable, but not erratic enough to cause any concern. If SurveyUSA's data were significantly influenced by children picking up the phone and mashing keys, we would not observe this stability.

Table 7-2: Pairwise Correlations with November 2006 Data

Month	Pro-Life	Pro-Choice	Republican	Democrat	Independent
Sep	0.94	0.94	0.90	0.88	0.91
Jul	0.94	0.94	0.93	0.90	0.95
May	0.94	0.92	0.89	0.86	0.91
Mar	0.94	0.94	0.91	0.86	0.93
Jan	0.95	0.94	0.90	0.88	0.92

Critics have raised important conceptual objections to the IVR methodology. But these data suggest no reason to question SurveyUSA's data. The onus is on the critics to produce empirical evidence that these data should not be trusted.

Works Cited

Ansolabehere, Stephen, and James M. Snyder, Jr. 2002. The incumbency advantage in U.S. elections: An analysis of state and federal offices, 1942-2000. *Election Law Journal* 1 (3): 315-338.

Atkeson, Lonna Rae, and Randall W. Partin. 1995. Economic and referendum voting: A comparison of gubernatorial and senatorial elections. *American Political Science Review* 89 (March): 99-107.

Atkeson, Lonna Rae, and Randall W. Partin. 1998. Economic and referendum voting and the problem of data choice: A reply. *American Journal of Political Science* 42 (July): 1003-1007.

Atkeson, Lonna Rae, and Randall W. Partin. 2001. Candidate Advertisements, media coverage, and citizen attitudes: Agendas and roles of Senators and governors. *Political Research Quarterly* 54 (December): 795-813.

Black, Gordon. 1972. A theory of political ambition: Career choices and the role of structural incentives. *American Political Science Review* 66 (March): 144-159.

Bloom, Joel David. 2003. Reliable compared to what? Empirical tests of the accuracy of election polls, 2002. Expanded version of a paper presented at the Annual Meeting of the American Association of Public Opinion Research, May 2003, in Nashville, Tennessee.

Bloom, Joel David, and Jennie Elizabeth Pearson. 2005. A probability-theory based test of the reliability of election polls. Paper prepared for presentation at the 2005 Annual Meeting of the American Association of Public Opinion Research, May 2005, in Miami, Florida.

Brace, Paul. 1993. *State government and economic performance*. Baltimore, MD: Johns Hopkins University Press.

Carsey, Thomas M., and Gerald C. Wright. 1998. State and national factors in gubernatorial and senate elections. *American Journal of Political Science* 42 (July): 994-1002.

Chubb, John. 1988. Institutions, the economy, and the dynamics of state elections. *American Political Science Review* 82 (March): 133-154.

Crew, Robert E. Jr., and Gregory R. Weiher. 1996. Gubernatorial popularity in three states: A preliminary model. *Social Science Journal* 33 (1): 39-54.

Ebeid, Michael, and Jonathan Rodden. 2006. Economic geography and economic voting: Evidence from the US states. *British Journal of Political Science* 36: 527-547.

Erikson, Robert S., Gerald C. Wright, and John P. McIver. 1993. *Statehouse democracy: Public opinion and policy in the American states*. Cambridge: Cambridge University Press.

Fiorina, Morris. 1978. Economic retrospective voting in American national elections: A micro-analysis. *American Journal of Political Science* 72: 426-443.

Hibbing, John R., and John R. Alford. 1981. The electoral impact of economic conditions: Who is held responsible? *American Journal of Political Science* 25: 423-39.

Holbrook, Allyson L., Alison Pfent, and Jon A. Krosnick. 2003. Response rates in recent surveys conducted by non-profits and commercial survey agencies and the news media. Paper presented at the American Association for Public Opinion Research Annual Meeting, Nashville, Tennessee.

Holbrook-Provow, Thomas. 1987. National Factors in Gubernatorial Elections. *American Politics Quarterly* 15: 471–83.

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.

Jacobson, Gary C., and Samuel Kernell. 1983. *Strategy and choice in Congressional elections*. New Haven: Yale University Press.

Kenner, David, and William Saletan. 2004. Who nailed the election results? Automated pollsters. *Slate* (December 9), at http://www.slate.com/id/2110860/.

Kenney, Patrick J. 1983. The effects of state economic conditions on the vote for governor. *Social Science Quarterly* 64 (March): 154-161.

Kernell, Samuel. 1977. Presidential popularity and negative voting: An alternative explanation of the midterm Congressional decline of the president's party. *American Political Science Review* 71: 44-66.

Kramer, Gerald. 1971. Short-term fluctuations in U.S. voting behavior, 1896-1964. *American Political Science Review* 65 (March): 131-143.

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

Leyden, Kevin, and Stephen Borrelli. 1995. The effect of state economic conditions on gubernatorial elections: Does unified government make a difference? *Political Research Quarterly* 48 (June): 275-290.

Long, J. Scott, and Jeremy Freese. 2006. *Regression models for categorical dependent variables using Stata*, 2d edition. College Station: Stata Press.

Niemi, Richard G., Harold W. Stanley, and Ronald J. Vogel. 1995. State economies and state taxes: Do voters hold governors accountable? *American Journal of Political Science* 39 (November): 936-957.

Peltzman, Sam. 1987. Economic conditions and gubernatorial elections. *American Economic Review* 77 (2): 293-297.

Raftery, Adrian E. 1995. Bayesian model selection in social research. *Sociological Methodology* 25: 111-163.

Riker, William H., and Peter C. Ordeshook. 1968. A theory of the calculus of voting. *American Political Science Review* 62 (March): 25-42.

Rusk, Jerrold G. 2001. *A statistical history of the American electorate*. Washington, DC: CQ Press.

Simon, Dennis M. 1989. Presidents, governors, and electoral accountability. *Journal of politics* 51 (May): 286-304.

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

Stein, Robert M. 1990. Economic voting for governor and U.S. senator: The electoral consequences of federalism. *Journal of Politics* 52 (February): 29-53.

Tompkins, Mark E. 1984. Have gubernatorial elections become more distinctive contests? *Journal of Politics* 50 (Feb): 192-205.

Traugott, Michael. 1995. “CATS” dissected. *Public Perspective* 6 (August/September): 1-3.

Tufte, Edward R. 1975. Determinants of the outcomes of midterm Congressional elections. *American Political Science Review* 69: 812-26.

Tufte, Edward R. 1978. *Political control of the economy*. Princeton: Princeton University Press.

Turett, Stephen. 1971. The vulnerability of American governors, 1900-1969. *Midwest Journal of Political Science* 15 (Feb): 108-132.

Weatherford, M. Stephen. 1978. Economic conditions and electoral outcomes: Class differences in response to recession. *American Journal of Political Science* 22: 917-938.