Appendix

A: Robustness Checks Addressing Potential Strategic Donation

As noted in the main text, a potential concern about CFscores is that they are sensitive to the volume of receipts attracted by legislators, or by other irregulaties in donation behavior. For example, while some important legislators may attract donations from both ideological and access-seeking givers, others may enjoy donations from only one of these types. If a legislator receives donations only from access-seekers while another receives donations from ideological and access-seeking interests, CFscores may erroneously measure the former legislator as more moderate than the latter legislator, for example. This concern is particularly salient when examining state legislatures, where the volume of donations is typically lower than in Congress.

To address this concern, I introduce a variable, *Total Receipts*, into the regression models from Test 2. This variable represents the total dollar amount of donations received by the winning candidate. Its inclusion controls for the possibility that the total number of campaign funds received by the winner explains the observed correlation between winner moderation and exposure to same-party general-election competition. Table A1 depicts the results of regression models including this variable.

As Table A1 depicts, the inclusion of *Total Receipts* does not alter the substantive results presented in the main text. That is, even after controlling for the volume of campaign funds received by the winning candidate, exposure to same-party general-election competition remains negatively associated with ideological extremism among the winning candidates. *Total Receipts* is itself moderately associated with the ideological extremism of candidates, displaying a positive, statistically significant association with *Winner Extremism*. This relationship is only significant, however, when state-level fixed effects are included. This is perhaps not surprising, given differences in fundraising between California and Washington.

Table A1: Same-Party Competition, Campaign Receipts, and Winner Extremism

	Dependent variable:			
	Winner Extremism			
	(1)	(2)	(3)	(4)
Same-Party Competition	−0.144***	−0.148***	−0.095**	−0.097**
	(0.043)	(0.043)	(0.043)	(0.043)
Total Receipts	0.000	0.000	0.00000**	0.00000*
	(0.00000)	(0.00000)	(0.00000)	(0.00000)
Congressional Race	−0.028	0.046	0.014	0.016
	(0.071)	(0.082)	(0.069)	(0.079)
Party	−0.0003*	−0.0002	−0.0003	−0.0003
	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Incumbent Winner	−0.067	−0.077	−0.103*	−0.098*
	(0.055)	(0.056)	(0.053)	(0.053)
Open Seat	−0.020	−0.033	−0.034	−0.040
	(0.058)	(0.059)	(0.056)	(0.056)
Upper Chamber	−0.062	−0.058	−0.071*	−0.070*
	(0.038)	(0.037)	(0.036)	(0.036)
District Extremism	0.213***	0.236***	0.285***	0.270***
	(0.068)	(0.071)	(0.067)	(0.068)
Difference in Extremism	−0.357***	−0.353***	−0.367***	−0.363***
	(0.038)	(0.038)	(0.036)	(0.036)
Constant	1.115***	1.071***	1.143***	1.083***
	(0.063)	(0.070)	(0.061)	(0.067)
State FEs?	N	N	Y	Y
Year FEs?	N	Y	N	Y
Observations	324	324	324	324
Log Likelihood	−24.028	−21.190	−11.248	−6.433
Akaike Inf. Crit.	68.056	68.379	44.495	40.867

Note: *p<0.1; **p<0.05; ***p<0.01

Beyond concerns associated with the volume of donations, one may still reasonably worry that the unusual electoral environment established by the top-two primary may encourage differential contribution patterns accross same-party and two-party general-election contests. More specifically, the removal of inter-partisan competition disallows donors both inside and outside the state/district from simply donating to their copartisans.[1] Thus, one may argue that tests of the paper's main results should be re-examined, using roll-call or position-taking based measures alone.

Although concerns about differential donation patterns are important to consider, the main text focuses specifically on CFscores for several key reasons. First and foremost, the focus of the theory underlying the paper's hypotheses lies in the electoral arena: to what extent do candidates exposed to same-party, general-election competition present themselves as the more moderate of the two general-election candidate? To be sure, revealed ideology within the legislature is ultimately a central concern for top-two primary proponents (and opponents); but, given that such revealed preferences are themselves potentially influenced by agenda-setting behavior, it is useful to first consider whether members present themselves more or less ideologically to donors and the public during the campaign. Indeed, inasmuch as the goal of partisan agenda-setting within the legislature is to present as unified a front as possible, examining electoral behavior provides an opportunity to examine whether candidates appear to present themselves as a more or less moderate candidate for office.

In addition to theoretical reasons for using electoral-focused rather than legislature-focused measures, CFscores provide a variety of practical advantages over scores such as Shor and McCarty's (2011) NPAT scores, given the study's race-level (and not legislator-level) level of analysis. Perhaps most crucially, CFscores exist for both winning and losing candidates, while measures based on roll calls alone do not. This fact enables the measurement of two variables featured in the main text—one dependent variable and one

[1] I thank an anonymous reviewer for articulated this point especially well.

independent variable. First, by allowing one to compare the ideological extremity of general election opponents, CFscores enable tests involving the outcome variable, *Election of Moderate*. Without a measure of the losing candidate's preference, this entire set of tests would be rendered impossible. Second, CFscores enable the measurement of the variable *Difference in Extremism*, which captures how much more ideologically extreme general-election candidates are from one another. Inclusion of this independent variable is especially crucial, as it is a substantively and statistically predictor of the election of moderate candidates throughout all three sets of tests in the paper—demonstrating that, if the gap between two candidates is sufficiently wide, voters may possess the ability to discern which is more or less moderate. Finally, unlike Shor-McCarty scores, CFscores place both state legislators and members of Congress onto a single ideological scale. Given that the top-two primary applies to both state- and federal-level elections in California and Washington, relying primarly on roll-call scores like Shor-McCarty would necessitate removing federal races from the data. Given that federal cases represent some of the most high-profile instances of same-party competition to date, removing these races would render the paper's tests only a partial examination of the hypotheses of interest.

Still, in spite of these reasons for focusing primarily on electoral behavior, one would still hope that the main findings presented in the paper stand up to robustness checks that relate more directly to the legislative process. As I demonstrate in Table A2 below, the these text do in fact persist when measuring extremism using preference measures other than CFscores.

In the table, I retest models of the *Winner Extremism* outcome variable, using an alternative measure of preferences, DW-DIME (Bonica 2018), more closely tied to the legislative process. To place candidates into a single ideological space, DW-DIME uses information from both campaign donations and roll-call behavior. To do so, the scores' estimation procedure uses a sophisticated supervised machine learning technique to map donation patterns to a target measure of legislative voting behavior. In the case of

members of Congress, the target measure of voting behavior is DW-NOMINATE; for state legislatures, it is Shor-McCarty. Impressively, Bonica's tests show that this approach correctly classifies actual roll call votes at a rate nearly identical to DW-NOMINATE for all members of Congress, and at a *better* rate than DW-NOMINATE for first-time members. Thus, the scores are both firmly tied methodologically to actual legislative behavior and predict future legislative behavior at high rate, as they derive their latent ideological structure not solely donation patterns but from the roll call measures from which the model learns. Additionally, the method generates ideology estimates for losing candidates, should one wish to replicate analysis of the *Election of Moderate* outcome variable.

Using this measure, I find similarly strong support for the notion that winners of same-party general-election are more moderate than simiarly situated winners exposed to traditional two-party competition. These results are summarized in Table A2. As the table summarizes, exposure to same-party general election competition is consistent associated with the election of a more moderate candidate overall, regardless if one introduces fixed effects at the state- or year- level. Taken together, then, this analysis and the donation volume analysis indicate that the paper's main results are not merely artifacts of unusual or overly strategic donation patterns.

Table A2: Re-Testing Winner Extremity Hypotheses with Roll-Call-Focused Measure

	Dependent variable:					
	|DW-DIME|					
	(1)	(2)	(3)	(4)	(5)	(6)
Same-Party Competition	−0.042***	−0.043***	−0.027**	−0.028**	−0.090***	−0.090***
	(0.012)	(0.013)	(0.013)	(0.013)	(0.030)	(0.030)
Post Reform		0.005				
		(0.009)				
Democrat	−0.0002***	−0.0002***	−0.0005***	−0.0005***	0.00000	−0.00002
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0002)	(0.0002)
Incumbent	−0.012	−0.027	0.002	0.004	−0.003	−0.005
	(0.017)	(0.017)	(0.018)	(0.019)	(0.039)	(0.040)
Open Seat	−0.015	−0.015	−0.021	−0.019	−0.018	−0.018
	(0.018)	(0.018)	(0.019)	(0.020)	(0.041)	(0.041)
State Upper Chamber	−0.013	−0.013	−0.004	−0.004	−0.019	−0.015
	(0.009)	(0.009)	(0.011)	(0.011)	(0.025)	(0.025)
District Extremity	0.070***	0.069*	0.041*	0.040*	0.109*	0.109*
	(0.018)	(0.018)	(0.023)	(0.024)	(0.055)	(0.056)
Difference in Extremity	−0.048***	−0.047***	−0.063***	−0.063***	−0.053**	−0.056**
	(0.009)	(0.009)	(0.012)	(0.012)	(0.024)	(0.024)
California	0.090***	0.092***	0.082***	0.072***	0.080*	0.073*
	(0.009)	(0.009)	(0.011)	(0.014)	(0.041)	(0.041)
Washington					−0.045*	−0.032
					(0.026)	(0.028)
Constant	0.368***	0.365***	0.416***	0.399***	0.403***	0.392***
	(0.018)	(0.019)	(0.026)	(0.028)	(0.059)	(0.063)
State-Level Effects	Y	Y	Y	Y	Y	Y
Year-Level Effects	N	Y	N	Y	N	Y
Observations	407	407	281	281	129	129
Log Likelihood	478.186	478.341	334.861	337.132	101.866	103.915
Akaike Inf. Crit.	−938.373	−936.682	−651.722	−650.265	−183.731	−181.830

Note: *p<0.1; **p<0.05; ***p<0.01

B: Facsimile of Sam Reed Meeting Notes

In August 2016, the author visited the Washington State Archive in Olympia, WA to study the records and writings of the architects of the top-two system. The primary architect of the top-two primary system—and primary defender in front of the Supreme Court—was then-Secretary of State Sam Reed. Given transparency laws in place in Washington, most or all of Secretary Reed's e-mails, meeting notes, and other office contents from his time in office have been preserved at the Archive and are available for public viewing. Thus, as part of the trip, the author photocopied many of these documents, along with columns, new stories, and ballots from the nearly decade-long battle over the top-two. These included the clip found in Figure 1 of the main text.

Below, I display the full photocopy of the document from which Figure 1 was created. It should be noted that the sentiments captured in the notes were independently reiterated in a later phone interview with Secretary Reed directly.

What is it your constituents want?
Compliment caucus.
— Why? Broken?

This bill is good for Republicans. BOP Intervenor

- Libertarian Issue (Slade, other examples)
- Ds are larger party--we get hurt when Ds pick party ballot in primary
- We need cross over votes in many districts
- We need to elect moderates in Urban areas

This bill is the only vehicle with a realistic chance of passing.

- We have the votes to move the bill in the House with R help
- We doubt that Montana has the support to move in the house.
- Top two is the only bill with a realistic chance.

The Modified Montana has many problems (We got the support.)

- The bill has many technical concerns
- It would cost a lot to administer
- The Auditors and my staff are very concerned about administration
- The bill requires voters to choose a party ballot
- The bill controls ballot access for candidates — Statewide 350,000
- The bill gives Party Rules the Force and Effect of Law
- The bill gives the parties control over the election process
- The parties could change the rules, and election process, every year
- Parties could opt out of a primary and disenfranchise many voters
- Montana style bill puts us on slippery slope of party litigation

Please Don't believe everything you hear from the Parties.

- Top Two may result in litigation--but we will win.
- Parties may litigate right to use "name"--but we will win
- Parties may go to convention--but will not control ballot access
- Montana results in more "serious" litigation--and we might not win

Modified Blanket Primary has withstood the test of time and legislative process

- Public supports 2-1
- 34 editorials support

Booth
California
Norm Rice
Larry Phillips
Why 3/4 of BOP in Senate Against it: Adam Kline, Jeane Kohl-Wells, Jacobson, Rosa Franklin

C: Generating Matched Control Units Using the Population of Elections Available in DIME

In the main text, I present a matching analysis in Test 3 that compares same-party general-election contests in Washington and California to a random same of contests from non-top-two states across the U.S. As I indicate there, my primary reason for focusing on a relatively small random sample is practical: organizing a dataset that matches a large number of both winning candidates and their *main* general-election challenger is challenging. This is especially true when dealing with occasional missing data in Bonica's DIME dataset. Nevertheless, matching winning and losing candidates is important, because it allows the tests to match on key covariates like *Difference in Extremism*.

Nevertheless, particularly given the asymptotic properties of matching methods, there are methodological advantages to choosing the matched control set from the population of races available in the DIME dataset. First, doing so ensures excellent ballance between **X** covariates, since the matches are selected from a wide array of races. Second, particularly when calculating the ATC, drawing from the population of races generates a larger matched sample than in the random-sample case. Given these advantages, I execute a robustness check of my Test 3 results by generating my matches from the population of legislative races available in Bonica's DIME dataset. While doing so precludes me from matching on variables such as *Difference in Extremism*, I am able to match on whether the race's winner was an incumbent, whether the races was over an open seat, the chamber of the contest, the extremity of the district in question (folded Tausonovitch-Warshaw scores), party of the winner, and year of the contest.

The results of these matching analyses are summarized in Table A3 and are overall quite consistent with those presented in the main text. Indeed, races exposed to same-party general-election competition tend to elect more moderate candidates overall than do highly similar districts that were not exposed to such competition.

Table A3: Matching Analyses for Test 3, Using Population of Races

	(1)
Estimate (ATT)	-0.173
AI Standard Error	0.070
p-value	0.013
Original Number of Treated Obs.	54
Matched Number of Treated Obs.	54
Estimate (ATC)	-0.163
AI Standard Error	0.122
p-value	0.181
Original Number of Control Obs.	12741
Matched Number of Control Obs.	12741