Googling the Top Two: Information Search in California’s Top Two Primary

Betsy Sinclair
Washington University in St. Louis

Michael Wray
Washington University in St. Louis

Abstract

After California’s adoption of the top two primary, voters faced the possibility of ballot choices between co-partisan candidates (two Democrats, for example, or two Republicans). We use the publicly available Google Trends data, which provides the rate of searching for particular words, to evaluate whether Californians are more likely to search for the names of legislators who faced co-partisan challengers in their general election than to search for the names of legislators who faced opposite-partisan challengers in the general election. We find evidence of increased search for the general election and, moreover, find that there is no increase for the primary election, suggesting that when the typical voter loses a key electoral cue (the party label) the voter will rely upon other sources of information to make a voting decision.

1 Corresponding author. Associate Professor of Political Science, Washington University in St Louis, 1 Brookings Drive, St Louis, MO 63130. Email: bsinclai@wustl.edu.
Googling the Top Two:  
Information Search in California’s Top Two Primary

Betsy Sinclair  
*Washington University in St. Louis*

Michael Wray  
*Washington University in St. Louis*

Introduction

Voters in the 76th state assembly district in California faced a difficult choice for representative in the November 2012 general election. Their general election ballot included the names of two co-partisan Republicans candidates, Rocky Chavez and Sherry Hodges. Not only were voters unable to rely upon party label distinctions between the candidates as a cue for voting, but also the candidates were relatively similar in terms of their visibility and publicly stated platforms. Neither candidate had any state-level electoral or legislative experience.

Both candidates positioned themselves as fiscal conservatives who, once in office, would focus on job creation and fiscal reform. As fellow Republicans, they agreed on drilling for oil, repealing the DREAM Act, rejecting tax increases, and a number of other social and political issues. While the candidates differed on a number of more nuanced points (Chavez focused on veterans issues while Hodges focused on economic and education reform) their political positions were remarkably similar. With the little differentiation and the same party label, how would voters decide for whom to vote?

We suggest the possibility that voters bear the burden of participatory democracy in this context by seeking out additional political information online. During the months leading up to the general election, Google searches for Chavez and Hodges increased by thirty and fifty percent, respectively. If candidates from the same party run against each other, voters encounter a political environment with more nuanced differences between candidates that cannot be simplified by party labels. How do voters decide for whom to vote in the absence of partisan cues?

---

2 Corresponding author. Associate Professor of Political Science, Washington University in St Louis, 1 Brookings Drive, St Louis, MO 63130. Email: bsinclai@wustl.edu.

3 [http://www.smartvoter.org/2012/06/05/ca/state/race/caasm76/](http://www.smartvoter.org/2012/06/05/ca/state/race/caasm76/)

4 See Figure 1. We observe a sharp spike in the Google searches for both candidates ahead of the general election.
Figure 1. Chavez and Hodges 2012 Search Volume

![Search Over Time](image)

Note: This figure indicates the Google search volume for Republican assembly candidates Rocky Chavez and Sherry Hodges. As is visible from this figure, there was a sharp spike in searching in California associated with the November 2012 general election.

California voters face this very problem. In June 2010, California voters passed Proposition 14, which shifted the state from a partisan primary system in which one candidate from each party would face each other in the general election to a non-partisan “top-two” primary format. Within the new primary format, all the candidates for the congressional and state elective offices of all the parties are listed on the primary ballot. All voters then cast votes, regardless of party alignment or affiliation. The two candidates who receive the highest number of votes within the primary then move on to face each other in the general election, even if they have the same party label.⁵

The new primary system, inevitably, periodically leads to candidates of the same party campaigning against one another in the general election. In the general election in November 2012, for example, there were more than 20 contests from the State Assembly and State Senate that contained two candidates from the same party (Kousser et al 2014).⁶ These new co-partisan contests provide an opportunity to understand how the changes in the primary system and the subsequent absence of differentiation in party labels shift the public’s role in gaining political information.

---

⁵ [http://voterguide.sos.ca.gov/past/2010/primary/propositions/14/analysis.htm](http://voterguide.sos.ca.gov/past/2010/primary/propositions/14/analysis.htm)

⁶ See Figure 2 for a map of California that indicates the location of the 28 contests that had a co-partisan challenger in the general November 2012 election.
Political information is necessary for citizens to make informed voting decisions within a democratic process. Political scientists have long theorized on how electoral information is gained, absorbed, and disseminated. Central to any information environment and structure during election season, though, are political parties. Parties are responsible for backing a candidate, funding public engagement, and informing the citizens (Masket 2011). Party labels provide voters with low-cost information to identify if a particular candidate agrees with their ideology, policy, and viewpoints. What happens, then, to the information environment when parties can no longer play a central role in informing the public?

When citizens can no longer rely on party labels to make political decisions, they are forced to rely on other information outlets. The Internet provides voters with the opportunity to strategically seek out political information most relevant to them. Hall, Sinclair, and Sinclair (2014) argue that citizens engaged in online searches ahead of the U.S. presidential primary elections in 2008 to cast more informed votes. In these primaries, individuals were forced to choose between candidates within the same party. We argue that the pressures of this choice are similar for some voters in the 2012 California election cycle. In particular, those voters whose districts present them with same-party general election candidates without the convenience of party labels are confronted with a more nuanced election in which party identity fails to provide low-cost cues for voters. Records of Internet searches conveniently provide the unique opportunity to quantify the public’s search for political knowledge in these low-information environments.

In the section that follows, we argue that people will be more likely to seek out political information online when candidates face a co-partisan challenger in the general election. We then turn to a series of empirical analyses. By comparing online searches for legislative winners who either faced or did not face a co-partisan general election challenger, we can begin to
understand if and when citizens seek out additional information relevant to their votes after California’s primary reform. A difference in search volume when candidates face co-partisan challengers as opposed to candidates from opposing parties suggests that the new “top-two” primary generates a low-information environment in which voters seek out relevant information directly. We compare differences across searches to discern whether search patterns vary by differences in legislators or, rather, differences associated with the absence of partisan general election cues. We conclude with thoughts on the implications of increased searching for whether the top two primary election reform will generate more moderate and pragmatic legislative representation.

**Elections and Political Information**

For the most part, voters have no need to possess comprehensive political knowledge. During more salient moments like elections, however, the average voter is incentivized to gain sufficient information to cast a vote. In these moments, voters do not rely exclusively on their own knowledge, instead deferring to the information provided by political parties, interest groups, political figures, and other cues (Lupia and McCubbins 1998). These sources do not create a comprehensive understanding of the political landscape; they do provide key information to voters as to who is considered a viable candidate. By gathering new information about the candidates and other votes’ choice, citizens can vote strategically – that is, they can procure enough information about what other voters are likely to do so as to not ‘waste’ their votes on candidates that are unlikely to be elected (Issacharoff 2004). Other low-cost informational cues provide voters with the necessary knowledge to cast an intelligent and strategic vote.

Party primaries give voters insight into who are viable candidates. Once a primary has selected a particular candidate, the voter’s decision is simplified as the voters can rely upon the party label to summarize a candidate’s ideological platform (Jessee 2010; Bartels 2000). Thus, the party label is a key characteristic when voters decide for whom to cast their ballot. Without the party label, voters are forced to operate in a lower-information environment in which they must seek alternative information in order to cast an informed vote that is not wasted on a non-viable candidate.

Without party labels to provide cues as to whom to vote for, the voter seeks out cues from alternative sources. Some political scientists contend that descriptive characteristics, including age, gender, or race, give voters low-cost information about the candidates and give voters insight into where to most strategically cast their ballot (Campbell 1983; Geer 1989; Kenney and Rice 1992; Marshall 1981, 1984; Norrander 1986; Pfau et al. 1993, 1995; Stone, Rapoport, and Abramowitz 1992; Williams et al. 1976). More recent literature has suggested that with the emergence of multiple information channels, voters are turning increasingly to the Internet to search for political information in order to cast a strategic vote (Mossberger, Tolbert, and McNeal 2008; Best and Krueger 2005; Kenski and Stroud 2006; Drew and Weaver 2006; Dalrymple and Scheufele 2007; Hall and Sinclair 2011).

The Internet provides voters with a low-cost opportunity to marginally increase political knowledge in order to cast a more informed vote. Voters use the Internet to examine candidates’ platforms and fulfillment of campaign promises (Anderson and Cornfield 2003; Bimber 2003; Conrnfield 2004). Within the 2012 campaign, 47% of the adult population used the Internet to
gain political information (growing from 44% in 2008). As voters turn increasingly to online resources to understand more about the candidates, we can glimpse into what prompts voters to seek out more information about the candidates.

The California primary reform provides an excellent opportunity to study the role that information supply has in elections. Before the reform, California elections were almost exclusively partisan, positioning one Democrat against one Republican. Candidates, funded by parties, had some visibility through traditional party mechanisms. The reform allows us to compare those races in which there are no co-partisan challengers with elections with a co-partisan runoff. By comparing two legislators’ online search volume, we can gauge the public’s willingness to seek out political information when deciding to cast their vote in the absence of party labels.

We are thus focused on the relationship between search volume and legislative contests where there was a co-partisan challenger in elections that occur after the implementation of the top two primary election reform. We have two possible hypotheses regarding an increase in search volume associated with a co-partisan challenger. First, an increase in search volume indicates a lower-information environment in which parties no longer have a central role in disseminating information. If citizens are using the Internet to seek out more political information, the new primary system has given nontraditional sources a more central role in informing the public.

The second possible explanation for an increase in search volume is a more competitive election. One of the goals of the California election rule reform was to create more competitive state elections. Proponents of the top-two primary system contended that through the adoption of the system, the incumbents would not have a safe seat, leading to an increase in electoral competition. Some scholars have found that the reforms have in fact made legislative races more competitive than in recent years (McGhee and Krimm 2012).

We next describe our data sources and present our strategies for operationalizing a test of our key hypotheses, that searching increases in general elections where co-partisans compete.

**Google Trends Data and Co-Partisan Challengers**

With the large variety of information outlets and channels through which citizens can gain information, it is difficult to capture political interest over time. With the Internet, we can now use search engine data to gain some insight into how and when citizens are directly seeking out political information. The search engine Google, in particular, keeps a public record of search queries through “Google Trends.” We focus on Google Trends records in California in the electoral era following the implementation of the top two primary.

In order to understand the effect of the adoption of the top two primary on the search for political information, we examine search data provided by the “Google Trends” application. The application captures public interest through Google search data, which is normalized on a scale of 0-100 to reflect the number of searches that have been done for a particular term relative to the total number of searches done on Google over time. We used this application to aggregate search volume data for particular legislators from June 2010 to February 2013 within California.

---

8 https://support.google.com/trends/answer/4355164?hl=en
We gathered search volume data for particular California legislators by tracking the search queries associated with their names.

For example, we gathered data on the search term “brian jones” corresponding to the San Diego Assembly representative Brian Jones. When someone inquiring about Brian Jones searches his name, Google archives the search and accumulates the number of queries. Google then compares the number of search queries that the term “brian jones” has received within a particular time period to the relative total number of searches on Google done over time, ultimately calculating a number to indicate the popularity of that search term. For Brian Jones, the search volume varies from a rating of 76, indicating high search volume during November 2012, to a rating of 22, indicating lower search volume during September 2010. The data effectively captures citizen’s search for information over time. We use this data to compare search volumes between legislators who face co-partisan challengers and those who do not.

In particular, we collect the average search volume for each legislator who served in the California State Assembly or California State Senate following the November 2012 general election. We compare search volume in two periods: the period between the primary election (June 5) and the general election (November 6) and the period starting at the beginning of the year (January 1) and going to the primary election (June 5). Our goal is to test whether those legislators who faced co-partisan challengers have an increased volume of Google searches in both the primary election and the general election. We hypothesize that legislators are more likely to face increased searching in the general election (due to the absence of partisan cues) but are not necessarily likely to face increased searching in the primary election (as those voters who are likely primary voters are accustomed to voting in elections without the advantage of partisan cues).

In total, we gathered Google Trends data on 109 legislators, 72 members of the State Assembly and 37 members of the State Senate. We use this subset of members of the legislature in order to compare those legislators who are eligible to have faced co-partisan challengers. This population of individuals (all winners) allows us to also compare individuals who are similar to each other.

We also try to explain variation in search volume that may be associated with other factors. In particular, we control for the margin of victory of each legislator (to control for the competitiveness of the election) and whether each legislator serves in a leadership position as

---

9 The number of search queries limits the reliability and consistency of the Google Trends data. The first limitation is due to the amount of search volume certain legislators receive. Certain legislators do not have enough search volume to quantify a week-by-week scaled number. In these cases, we used their monthly search volume results, stretching the data over weekly periods in order to be consistent with other high-volume search terms. The second limitation is due to the limitations of the Google Trends application. The application only identifies search volume of certain terms if the number of queries reaches a particular volume threshold. If the term fails to be searched a certain number of times, the data will present a ‘zero’ for these periods of time. Our analysis is therefore limited in capacity due both to the weak interest in the search terms and the Google Trends application.

10 While there are 80 seats in the Assembly and 40 seats in the Senate, our data is constrained slightly by our need to associate these individuals with other covariates. We exclude all individuals who have similar names, thus reducing our sample size.
Table 1. Summary Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-primary Google Search Volume (2010)</td>
<td>3.33</td>
<td>0</td>
<td>50.37</td>
</tr>
<tr>
<td>Pre-primary Google Search Volume (2012)</td>
<td>10.21</td>
<td>0</td>
<td>63.11</td>
</tr>
<tr>
<td>Pre-general Google Search Volume (2010)</td>
<td>6.16</td>
<td>0</td>
<td>65.54</td>
</tr>
<tr>
<td>Pre-general Google Search volume (2012)</td>
<td>16.42</td>
<td>0</td>
<td>78.52</td>
</tr>
<tr>
<td>Co-partisans (1 indicates faced in a General Election)</td>
<td>.19</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Margin of victory</td>
<td>19.15</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>Legislative leader (1 indicates leader)</td>
<td>.15</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Note: These summary statistics are based on 109 observations from legislators who serve in the 2013-2014 legislative session. Google search dates for the pre-primary in 2010 run from January 1, 2010 until the primary election June 8, 2010 and for the pre-general in 2010 run from the primary election until November 2, 2010. Google search dates for the pre-primary in 2012 run from January 1, 2012 until the primary election June 5, 2012 and for the pre-general in 2012 from the primary election until November 6, 2012. 21 total contests had a co-partisan competing in the general November 2012 election.

This may affect the political media presence of the legislator. Finally, we explicitly control for the amount of Google searching that occurred during an identical search window in the previous election to account for those legislators who already had high public visibility. Table 1 presents summary statistics about our Google search variables, the number of legislators who faced a co-partisan challenger in the general election, and our other covariates.

Looking at the raw data it does appear that on average there is an increased search volume in the general election. Our analyses will discover whether that increase is associated more with particular kinds of contests – specifically, those with co-partisan challengers.

Results: Co-partisan Challengers Generate Increased Search

We analyze Google search data to look for evidence that voters are increasing their online searching when candidates face co-partisan challengers in the general election. Thus we focus on the Google searching that occurs prior to the legislative session immediately following implementation of the “top two” primary law (2013-2014). We hypothesize that we will observe an increased number of Google searches for those legislators who faced a co-partisan challenger in the general election. We focus on the searching patterns ahead of both the primary and general elections.

---

11 We consider the following leadership positions: for the Senate, we include President, President pro Tempore, Majority Leader, Minority Leader, Majority Whip, Minority Whip, Party Caucus Chairs, Secretary of the Senate, and Chief Sergeant-at-Arms. For the Assembly, we include Speaker, Speaker pro Tempore, Majority Floor Leader, Assistant Majority Floor Leader, Majority Whip, Assistant Majority Whip, Democratic Caucus Chair, Minority Floor Leader, Chief Clerk, and Chief Sergeant-at-Arms.
Table 2. Linear Regression Coefficients on Google Search Volume

<table>
<thead>
<tr>
<th></th>
<th>Pre-primary Search</th>
<th>Pre-primary Search</th>
<th>Pre-general Search</th>
<th>Pre-general Search</th>
</tr>
</thead>
<tbody>
<tr>
<td>Co-partisan Challenger</td>
<td>5.40</td>
<td>4.99</td>
<td>13.41*</td>
<td>15.16*</td>
</tr>
<tr>
<td></td>
<td>(3.66)</td>
<td>(3.32)</td>
<td>(4.51)</td>
<td>(4.27)</td>
</tr>
<tr>
<td>Leader</td>
<td>0.47</td>
<td>-0.31</td>
<td>0.45*</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>(3.52)</td>
<td>(4.56)</td>
<td>(.12)</td>
<td>(.10)</td>
</tr>
<tr>
<td>Previous Search Volume</td>
<td>0.68*</td>
<td>0.45*</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.13)</td>
<td>(.12)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Margin of Victory</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>9.17*</td>
<td>6.90*</td>
<td>13.84*</td>
<td>9.02*</td>
</tr>
<tr>
<td></td>
<td>(1.61)</td>
<td>(1.6)</td>
<td>(1.98)</td>
<td>(2.79)</td>
</tr>
<tr>
<td>Adjusted R2</td>
<td>.01</td>
<td>.19</td>
<td>.07</td>
<td>.20</td>
</tr>
<tr>
<td>N</td>
<td>109</td>
<td>109</td>
<td>109</td>
<td>109</td>
</tr>
</tbody>
</table>

Note: * indicates statistical significance using 95% confidence intervals.

Our primary hypothesis is that legislators who face co-partisan challengers will experience increased search volume in the general election. We anticipate this will not necessarily be true in the primary election; it is not clear whether voters will consider cross-over voting (voting for an opposite-party primary candidate) and would thus feel the need to research candidates from both parties or if instead they will rely upon their previous methods of information search. Below we produce the results from four separate linear regressions. In the first two, we evaluate whether having an eventual co-partisan challenger in the general election results in increased Google search volume ahead of the primary election. Thus our primary coefficient of interest is the coefficient on an indicator variable for whether the legislator faced a co-partisan challenger. In the first model we include no covariates, whereas in the second model we include the search volume from the previous election and whether the legislator served as a leader. In the third and fourth models, we evaluate whether having a co-partisan challenger in the general election results in increased Google search volume ahead of the general election. Here we anticipate that voters are unable to rely upon standard cues or habitual information search and will thus be more likely to search for information about the candidates. We produce the model again with and without covariates, adding one additional covariate to this last model: the margin of victory for the winner to control for competitiveness of the general election.

Looking at the coefficients from the primary election search period in the first two columns, it is clear that candidates who were (eventually) exposed to a co-partisan challenger experienced no statistically significant increase in search volume. Leadership also does not predict search volume, and unsurprisingly a high previous search volume is associated with a high current search volume. While it is difficult to draw conclusive inferences from the models from the primary, they do suggest that voters are unlikely to be engaged in high levels of additional...
searching despite the fact that the top two primary election law changes the kind of election faced in the district.

Yet, we observe a very different pattern when we focus on the pre-general election searching. Having a co-partisan challenger in the general election is associated with between 13-15 percent increase in Google (scaled) searches in the state of California ahead of the general election. These coefficients are statistically significant and substantively their magnitude indicates a sharp spike in searching is associated with facing a co-partisan challenger relative to facing a same-partisan challenger. In terms of the other covariates in the model, leadership does not significantly predict search volume but previous high search volume is associated with increased search. While previous search volume does have a statistically significant coefficient, however, the magnitude of this coefficient is actually quite small. Increasing the previous search volume by 10% would result in about a 6 percent increase in Google (scaled) searches in the state of California – about half of the effect we find relative to facing a co-partisan challenger. In terms of competitiveness, we find no statistically significant increase in searching associated with more competitive elections – that is, those elections with smaller margins of victory.

These results are consistent with voters who, faced with an election where they are unable to rely upon traditional party cues, need additional information to cast a vote. Interestingly we do not see a parallel increase that occurs ahead of the primary election. While this adds credibility to our results, it should not be interpreted to imply that voters are not also increasing their search habits ahead of the top two primary. It is simply the case that as all Californians participate in the top two primary we do not have as stark a comparison to draw from one set of legislators to another.

Conclusion

After California implemented the top two primary election reform, many voters were faced with two co-partisan choices in the general election in November 2012 for their state Assembly and Senate representatives. These voters were unable to rely upon party labels to distinguish the candidates. How, then, do the voters decide? This paper provides evidence that many California residents subsequently searched online for additional information.

Our empirical results demonstrate a clear difference in search volume between those legislators who faced a co-partisan challenger in their general elections and those who did not. While this association could be triggered by other factors, we attempt to control for those that are most likely – specifically, controlling for the competitiveness of the election, whether the legislator holds a leadership position, and the level of prior searching in the previous election. Each of these covariates could also have been associated with an increase in search volume, which would have the potential to confound the effect of competing against a co-partisan. Yet we find little evidence that these other covariates are generating the high increases in search; rather, our single greatest predictor of search volume is whether the legislator faced a co-partisan challenger. By pairing our two general election regressions (where we focus on the window between the primary and the general election) to our two primary election regressions (where we focus on the window ahead of the primary), we are able to isolate the spike in search volume as specifically occurring between the primary and the general election.

Increased searching in general elections where both candidates are co-partisans is likely a consequence of the implementation of the California top two primary election reform. Voters are unable to discern differences between the two candidates in terms of their party labels and
thus seek out additional information. In many ways, this is an encouraging sign from citizens in a participatory democracy, as it suggests that voters are in fact bearing their electoral burdens responsibly. Although our data does not tell us exactly who is searching, or what information the searchers find, it does very clearly show differences in search volume between these different types of general election contests.

Many reformers hoped that the implementation of the California top two primary would result in a new legislative body that espoused more moderation and pragmatism than previous legislative sessions. One of the principle complaints of reformers was that partisan control over the process to select candidates was such that voters were seldom given the opportunity to cast a ballot for a moderate candidate. By eliminating partisan labels in some contests where co-partisans compete in general elections, it seems likely that the reform has succeeded to the extent that it permits voters to seek out information to make their own decisions about candidates. Yet, electing moderate legislators requires that the underlying preferences of the voters be moderate themselves – otherwise this increased searching will simply result in the voters being able to carefully select extreme candidates.

As the California top two primary is a relatively recent reform, we have only a little data on which to make inferences about legislative and voter behavior. In our analysis we included twenty-one contests where legislators faced a co-partisan challenger out of the one hundred nine contests overall. As more elections are held, we will have the opportunity to analyze more data. Yet focusing on this early data provides a unique window into what voters will do when faced with an electoral environment with limited information. As more time passes after the reforms, parties are likely to develop other informational strategies to reach voters to provide them with cues during co-partisan elections. For now, we find that Californians will Google more as a consequence of the top two primary.

References