Primary Elections and Political Accountability:
What Happens to Incumbents in Scandals?\textsuperscript{1}

Shigeo Hirano
Department of Political Science
Columbia University

James M. Snyder, Jr.
Department of Government
Harvard University
and NBER

April, 2012

\textsuperscript{1}This paper is part of a project supported by National Science Foundation grants SES-0959200 and SES-0617556. The opinions, findings, and conclusions or recommendations expressed in this article do not necessarily reflect the views of the National Science Foundation. We thank Daniel Campos, Mason Fitch, Jasmine Senior, Nida Vidutis and Adam Zelizer for excellent research assistance.
Abstract

One key role of elections is to allow voters to remove politicians who perform poorly in office. We analyze the extent to which incumbents who are involved in relatively serious political scandals lose elections. More importantly, we assess the relative importance of primary and general elections in removing such incumbents. How often do incumbents involved in relatively serious scandals lose in the primary election? How often do they lose in the general election? How often is it the case that the primary election was probably necessary in order to remove the incumbent – i.e. would the incumbents in “safe districts” have been re-elected in the general election if they did not lose in the primary? We find that that incumbents in scandals are more likely to face a serious primary challenger compared to other incumbents. This relationship is even stronger when the incumbent represents a “safe district” – i.e. a district where she would probably have won the general election. Our estimates suggest that primary elections have an important role in removing incumbents in otherwise “safe districts.”
1. Introduction

Are primaries sometimes a good substitute for general elections? Based on raw counts of the number of competitive races, the answer would appear to be “no.” Even a strong defender of primaries would probably be forced to concede “only in open seats.”

This is not a new state of affairs. More than fifty years ago Turner (1953, p. 210) wrote, “The comparative usefulness of the primary as a method for selecting successors for retiring incumbents does not offset the fact that the primary is not a successful alternative to two-party competition in most parts of the United States.” Since then, the situation has, if anything, deteriorated even further. As Ansolabehere et al. (2006, 2007, 2010) document, the level of competition faced by incumbents in primaries fell steadily throughout the post-WWII period, and the incumbency advantage in primaries also increased greatly. Thus, for example, in the most recent decade, only 22% of incumbents in the U.S. House faced more than a token challenge in the primary.\(^1\)

Of course, it is possible that few incumbents deserve to be strongly challenged in their parties’ primaries. Why should we think that, under normal circumstances, there are candidates available to a party who clearly dominate the incumbent? Incumbents have experience and seniority, and track records that prove they can win elections. Given the large incumbency advantage that exists in general elections, a party would be reluctant, especially in competitive districts, to give up this electoral advantage unless the alternative is superior on some dimension of value to the party.\(^2\) Such challengers may be rare.

Sometimes, however, it seems clear that an incumbent deserves to be replaced. This is especially true of incumbents who have been involved in malfeasance of some sort – illegal activities, or activities that, while not illegal, violate the norms of proper behavior to an extreme degree. These are relatively rare cases, but they are especially useful in estimating a “lower bound” on the value of elections. Do elections at least help citizens remove the worst types of politicians from office?

The existing literature concludes that primary elections do little to remove incumbents, even in cases where we expect primary electorates to punish them. Several papers examine

---

\(^1\) Here, we define a token challenger as one who received less than 1% of the primary vote.

\(^2\) See Hirano, et al. (2009) for a model that incorporates this logic.
the effect of political scandals on primary election outcomes involving incumbents. These papers typically find that scandals hurt incumbents in terms of vote share. However, since the degree of competition in incumbent-contested primaries is so minimal on average, the typical conclusion drawn from these studies is that primaries matter little or not at all. For example, Brown (2006b, pages 8-9) states, “The scandal variables are statistically significant and large in magnitude in nearly all of the regressions. As predicted, incumbents are hurt by scandal and in the pooled regressions they are harmed more by morality scandals than by monetary scandals. Since incumbents, however, typically win primary elections by larger margins than they do general elections, the impact of losing more percentage points in a primary may have little or no effect on the outcome (the constants in these models suggest as much).” Some studies even find small and statistically insignificant effects. Welch and Hibbing (1997, page 233) note that for the 1980’s, “Primary defeats are even rarer than voluntary departures... More importantly, [the primary defeat rate] is only slightly higher than the primary defeat rate for incumbents not charged with corruption, despite the fact that an incumbent not charged with corruption almost never loses in the primary.”

Jacobson and Dimock (1994) do find some evidence that the representatives who were more seriously involved in the 1992 U.S. House banking scandal were also more likely to lose in the primary. However, this article does not discuss whether the primary was necessary to remove the malfeasant incumbents. A large number of incumbents involved in this scandal also represented districts that had competitive general elections. Presumably many of these incumbents would have lost the general election if they had not lost in the primary. Thus, from this analysis alone, it is difficult to assess the importance of primary competition in holding representatives accountable.

In this research note we argue that the existing empirical literature underestimates the

---


4Lazarus (2008) describes his findings on senate and gubernatorial races as follows: “... the coefficient on scandal is statistically significant and in the predicted direction only for one type of challenger: amateur out-party challengers in gubernatorial election. Thus, it does not seem as though the presence of a scandal results in the entry of a significant number of serious challengers of either party.”

value of primaries in punishing malfeasant incumbents. Primaries are likely to be most useful in safe seats where the incumbent are not likely to be held accountable in the general election. The intuition is relatively straightforward. Since the primary winners will most likely win the general election for safe seats, we would expect scandal-plagued incumbents in safe seats to attract more, and possibly higher-quality, primary challengers. A relatively attractive set of challengers, in addition to the fact that the incumbent is less attractive due to the scandal, will lead many primary voters to favor one of the alternatives rather than the incumbent. Thus, the incumbent will tend to receive fewer votes, and lose with a higher probability.

Using data for U.S. House incumbents running for re-election between 1978 and 2008, we find that incumbents who are involved in scandals and represent districts that are safe for their party have: (i) a much higher chance of facing multiple challengers in the primary relative to ordinary incumbents; (ii) a much lower expected vote share; and (iii) a much higher chance of losing their primary elections. In particular, we find that 23% of the incumbents who were involved in relatively serious scandals and defeated in their re-election bids would probably have won in the general election had they not lost in the primary, and therefore the primary was probably necessary to remove them from office. Thus, sometimes, primary elections serve a valuable function. This happens in relatively few cases – only 2.8% of incumbents who represent safe districts are involved in scandals – but these are especially important cases.

2. Scandals and Electoral Competition, 1978 to 2008

We analyze the impact of scandals on primary and general election competition over three decades, 1978 to 2008.

---

6Suppose, for example, that primary challengers have a 10% chance of defeating an incumbent. If the challenger makes it to the general election, they have a 50% chance of winning competitive districts and a 100% chance of winning safe seats. Now suppose an incumbent involved in a scandal increases the challengers chance of winning the primary to 50%. Then the overall expected probability of winning the election increases 20% for competitive seats and 40% for safe seats.

7Of course, we do not know whether these incumbents would also have lost in conventions or “smoke-filled rooms,” so we cannot assert that primaries are superior to other nomination methods.

We examine seven dependent variables. The first, denoted *Primary Vote Share*, is the incumbent’s primary election vote share in 1992, where uncontested incumbents are assigned 100% of the vote. The second, *# Primary Opponents*, is the number of primary election challengers the incumbent faced in 1992. The third, *Primary Uncontested*, is a dummy variable indicating that the incumbent was uncontested in the primary in 1992.\(^9\) The fourth, *Quality Challenger*, is a dummy variable indicating whether the incumbent was challenged by a “high quality” opponent in the primary.\(^10\) The fifth, *Won Primary*, is a dummy variable indicating that the incumbent won his or her primary in 1992. The sixth, *General Vote Share* is the incumbent’s general election vote share in 1992, where uncontested incumbents are again assigned 100% of the vote. The last, *Won General*, is a dummy variable indicating that the incumbent won the general election in 1992.\(^11\)

For the four primary election dependent variables, we include the lagged values of *Primary Vote Share*, *# Primary Opponents*, *Primary Uncontested*, to control for level of contestation in the incumbent’s primary in previous elections. We also include a party dummy variable, *Republican* indicating that the incumbent is a Republican; and dummy variable, *Freshman*, indicating that the incumbent is a freshman, i.e., that he was not an incumbent in his or her previous race, and *Special*, indicating that the incumbent won his or her seat in a special election and therefore is serving only part of a term. For the two general election dependent variable, we include three control variables to capture the partisan orientation of voters in each incumbent’s district: the lagged value of *General Vote Share*, the vote share won by the incumbent’s party in the previous presidential election, *Pres Vote Share*, and \((\text{Pres Vote Share})^2\)\(^2\) to capture the partisan orientation of voters in each incumbent’s district. We also include the *Republican*, *Freshman* and *Special* dummies. We normalize *Pres Vote Share* in

---

\(^9\)In defining *# Primary Opponents* and *Primary Uncontested*, we drop token challengers who received less than 1% of the vote.

\(^10\)We define quality as a candidate who has won an election for Congress, a major state level office, state legislative office, mayor, city council, or county legislative or executive office. The information regarding challenger quality were gathered from the ICPSR State Legislative data sets, newspaper reports, and various on-line databases, such as www.ourcampaigns.com.

\(^11\)We drop Louisiana from all of our analyses since it employs a unique, nonpartisan blanket primary and runoff system rather than ordinary primaries followed by a partisan general election. The primary election data used in this analysis is described in Ansolabehere, et. al. (2006, 2010) and Hirano et. al. (2010).
each year by subtracting the year mean.¹²

The main independent variable of interest is a dummy variable, Scandal, which is 1 if the incumbent was involved in a “serious” scandal and 0 otherwise.

To determine whether an incumbent was involved in a scandal as objectively as possible, we define Scandal = 1 only in the following cases: (1) the House Committee on Standards of Official Conduct investigated the incumbent; or (2) the Department of Justice investigated the incumbent; or (3) Congressional Quarterly Weekly Reports specifically discussed the scandal in at least one article. The sources for these are the Historical Summary of Conduct Cases in the House of Representatives (2004), and the Summary of Activities reports published after each Congress by the House Committee on Standards of Official Conduct; the Report to Congress on the Activities and Operations of the Public Integrity Section published annually by the U.S. Department of Justice; and Congressional Quarterly Weekly Reports. Of course, many scandals appear in two or more sources.¹³ As noted in the introduction, relatively few incumbents are involved in serious scandals. In our time period, about 2.8% of congressional incumbents running for re-election are involved in scandals.

In this analysis, we focus on incumbents involved in scandals who choose to run for re-election. We should note that more than one-third of incumbents involved in scandals retire without seeking re-election. Our analysis may underestimate the effect of scandals if the incumbents who retire would have done particularly poorly in either the primary or general election as a result of the scandal. Of course, this is true of previous analyses as well.

In a few cases it appears that a scandal “broke” – at least in terms of media coverage – between the primary and general elections. We searched online newspaper archives for articles published during the campaigns covering scandals involving incumbents, and found

---

¹²We also ran specifications for the primary election dependent variables that included Pres Vote Share or the lagged value of General Vote Share, but these variables were never statistically significant at the .05 level.

¹³We do not count cases where the House Committee on Standards of Official Conduct simply dismissed a complaint against the incumbent and Congressional Quarterly does not discuss it. In a few cases, the committee authorized but deferred its investigation at the request of the Department of Justice. These cases involved ongoing DOJ investigations we include them as scandals. We count all investigations by the Department of Justice as scandals, even if they did not ultimately lead to a conviction. We considered various ways to try to quantify the relative magnitudes of scandals, but we have not found an objective enough way to do this, especially for different scandals that occur at different times and under different circumstances. The mix of scandals is clearly heterogeneous.
12 cases where the media coverage jumped sharply after the primary election but before the general election. We do not distinguish these cases in the analyses below. If these cases are excluded, or if we assign them to the Scandal = 0 category for the primary election, then the estimated coefficients for the primary election dependent variables are generally 5-8% larger in magnitude than those reported below.

Basic Results

Our baseline specifications is a simple linear model including the various covariates discussed above. In addition to the baseline specification, we ran a variety of specifications to check for robustness. One of these is shown in Table 1. For the primary election dependent variables, we also ran specifications that include a polynomial function of time with the variables Year, Year\(^2\) and Year\(^3\) to capture the general decrease in primary election competition that has occurred over time.\(^{14}\) For the general election dependent variables, we also ran specifications that include party-by-year fixed effects, to capture national tides.

Table 1 presents the results. For all of the primary election dependent variables, the estimated coefficient on Scandal is substantively meaningful, and for three of the four it is statistically significant at the .05 level. In addition, the inclusion of the time trend variables or fixed effects have little effect on the estimates. Incumbents in relatively serious scandals received an approximately 16 percentage point drop in their vote shares relative to incumbents not involved in scandals. In addition, they faced about 0.77 more primary challengers on average, a 0.23 increase in the probability of facing a primary challenger, and a 0.22 increase in the probability of facing a “high quality” primary challenger. These incumbents also had about a 0.11 increase in their probability of losing their primary. To put these figures in context, only 26% of incumbents not in scandals were challenged in the primary, fewer than 1% lost their primary election, and the expected number of primary opponents for incumbents not in scandals was 0.36.

In the general election, incumbents in scandals had an approximately 10 percentage point decline in their vote shares compared to other candidates. Moreover, these malfeasant

\(^{14}\)The results are essentially unchanged if we include only the linear time trend or the linear and quadratic terms.
incumbents also observed a 0.20 drop in their probability of winning the general election. Only 4% of the incumbents not in scandals lost in the general election.

**Interactions with Seat Safety**

The analysis in the previous subsection provides an estimate of an average effect. In this section, we investigate whether the effect of a scandal on primary election outcomes varies depending on district partisanship. Interestingly, roughly half of the scandals are in competitive districts. If incumbents in competitive districts felt more electorally accountable to their constituents than those in safe districts, then we would have expected a disproportionate number of scandals involving incumbents representing safe seats.

We study the same five primary election variables as above. Now, however, the main independent variable of interest is an interaction term, $\text{Scandal} \times \text{Safe Seat}$. The Safe Seat variable is a dummy variable for whether the seat is above or below the median partisanship of incumbents’ districts, where partisanship is measured by the presidential vote share of the party of the incumbent in the district subtracting the year mean. The median is 56.5%.\(^\text{15}\)

Table 2 presents the main findings. For all four of the primary election dependent variables, the estimated coefficient on $\text{Scandal} \times \text{Safe Seat}$ is substantively meaningful and statistically significant at the .05 level. Since Safe Seat is a dummy variable, the coefficient on the Scandal variable itself is an estimate of the effect of a scandal in districts for seats which, on average, are expected to be competitive in the general elections. In all cases this estimate is statistically significant at the .05 level and indicates that scandals hurt the incumbent in the average district. Again, the inclusion of the time trend variables or fixed effects have little effect on the estimates.

The estimated coefficients on the interaction terms imply that scandals hurt incumbents in the primary election much more if they represent safe districts than if they represent marginal districts. Incumbents involved in scandals who represent safe districts lose their primary election 18% of the time. By contrast, incumbents involved in scandals who represent marginal districts do not observe a significant drop in their probability of winning their

\(^{15}\)The presidential vote data are from various issues of *The Almanac of American Politics* and *Politics in America*, and U.S. Census files.
Finally, although we do not present the results, we also ran regressions for the general election dependent variables with *Scandal*, *Safe Seat*, and the interaction term included as regressors. For these variables, the interaction term was always small and never statistically significant at the .05 level.

Most of the scandal-ridden incumbents from safe districts who lost in the primary would have won in the general election. The estimates in Table 1 suggest an expected vote share loss of about 10 percentage points (and the lack of a significant interaction term effect implies that the expected loss does not vary with seat safety as measured by the presidential vote). Adding back the incumbency advantage of 8%, and remembering that the presidential vote in these districts favors the incumbent by 60% or more, we estimate an expected vote percentage for these incumbents of 58% or more, suggesting a comfortable win for most. This is confirmed in our sample – only 6.25% of the incumbents from safe districts who were involved in scandals and won their primary elections lost in the general election. Thus, more than twice as many scandal-plagued incumbents from safe districts lost in the primary election.

3. Conclusion

The findings above are easily summarized. Incumbents involved in relatively serious scandals face a much higher probability of being challenged in the primary than ordinary incumbents, and a much higher chance of losing their primary elections. They are especially likely to be challenged in the primary, and especially likely to lose, when they are involved in scandals and represent districts that are “safe” for their party.

What do these results suggest about the value of primary elections? On one hand, even for incumbents involved in relatively serious scandals, the probability of losing in either the primary or the general election is well below 50%. On the other hand, it is unclear whether or not the observed probability is sub-optimal. Clearly this depends on the social costs of

\[16\text{In this analysis we do not condition on the event that the incumbent faces at least one primary challenger. When we restrict attention to the cases with at least one primary challenger the results are somewhat weaker. The results for *Primary Vote Share* and *# Primary Opponents* remains statistically significant, but the results for *Quality Challenger* and *Won Primary* are no longer statistically significant and the coefficient estimates are smaller in magnitude.}\]
allowing the incumbent involved in the scandal to stay in office. While estimating a quantity such as this is an intimidating task, it is essential for making progress in evaluating the overall benefits and costs of elections.

Our analysis may also underestimate the value of holding incumbents in safe districts accountable through primary elections, since incumbents in safe districts tend to be more senior and powerful than those from marginal districts. This is also true for incumbents involved in scandals. In our sample, the median number of terms served for incumbents involved in scandals from safe districts was 7, while the median number of terms served for those involved in scandals from non-safe districts was just 5. It is likely that party caucuses and conventions find it difficult not to re-nominate powerful members of the party such as senior members of congress, and this may give primaries an even more important role.

We may also be underestimating the role of primaries in holding incumbents accountable since some incumbents involved in scandals choose to retire rather than run for re-election, and they are excluded. Anecdotal evidence suggests that many of these incumbents would have faced serious primary opposition had they run. To take one example: in April 2007, the FBI raided the home of John Doolittle (R, CA) because of possible involvement by his wife in the investigation of Jack Abramoff; Doolittle had direct ties with Abramoff as well. In the summer of 2007, three Republicans announced their intentions to consider running against Doolittle in the next primary election. These included a city councilman and a state assemblyman, Ted Gaines, who stated “I think voters have lost faith in his [Doolittle’s] leadership ability... when you lose the moral ability to lead, you kind of have to re-evaluate.”

In addition, a prominent county party chairman announced that he would no longer support Doolittle. In January 2008, Doolittle announce that he would not seek re-election.

Furthermore, if primaries are a mechanism for holding incumbents in safe districts accountable, then incumbents should already be incorporating the potential loss of office in a primary election into their decisions about whether to engage in any type of political malfeasance. As we noted above, scandals were roughly equally likely to have occurred in a safe

\[17\] Peter Hecht, “GOP Assemblyman Gaines says he is planning to run for Doolittle’s House seat,” Sacramento Bee, August 30, 2007.
district as a competitive district. One interpretation of these descriptive statistics is that primary competition may provide as strong a deterrent as general elections for incumbents from engaging in malfeasance. Alternatively, this may also suggest that electoral considerations only weakly affect incumbents’ decisions to become involved in activities that could lead to a scandal. Whether electoral considerations can prevent incumbents from engaging in political malfeasance remains an open question.

Our findings indicate that primary elections can improve accountability by removing incumbents who face re-election in districts or regions dominated by one political party. This suggests that other electoral systems with similar characteristics may benefit from introducing primaries. In other countries, such as Italy, where some regions or municipalities are dominated by one party, primaries have become more widely used in subnational elections. Whether primaries have a similar role in holding incumbents accountable in other electoral systems is an open research question.
REFERENCES


### Table 1: Analysis of Scandals, 1978-2008

<table>
<thead>
<tr>
<th></th>
<th>Primary Vote Share</th>
<th># Primary Opponents</th>
<th>Primary Uncontested</th>
<th>Quality Challenger</th>
<th>Won Primary</th>
<th>General Vote Share</th>
<th>Won General</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>-0.158 (0.019)</td>
<td>0.769 (0.127)</td>
<td>-0.236 (0.038)</td>
<td>0.217 (0.034)</td>
<td>-0.108 (0.025)</td>
<td>-0.107 (0.011)</td>
<td>-0.196 (0.036)</td>
</tr>
<tr>
<td>with Year Polynomial</td>
<td>-0.157 (0.019)</td>
<td>0.761 (0.127)</td>
<td>-0.230 (0.038)</td>
<td>0.217 (0.034)</td>
<td>-0.108 (0.025)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>with Party-Year Fixed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Each cell presents the estimated coefficient on the dummy variable *Scandal* from a separate regression. Robust standard errors in parentheses. Number of observations in brackets.
Table 2: Scandals and District Partisanship, 1978-2008

<table>
<thead>
<tr>
<th></th>
<th>Primary Vote Share</th>
<th># Primary Opponents</th>
<th>Primary Uncontested</th>
<th>Quality Challenger</th>
<th>Won Primary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scandal × Safe Seat</td>
<td>-0.110 (0.037)</td>
<td>0.774 (0.242)</td>
<td>-0.089 (0.076)</td>
<td>0.116 (0.067)</td>
<td>-0.150 (0.048)</td>
</tr>
<tr>
<td>Scandal</td>
<td>-0.101 (0.022)</td>
<td>0.370 (0.116)</td>
<td>-0.189 (0.055)</td>
<td>0.156 (0.043)</td>
<td>-0.031 (0.022)</td>
</tr>
<tr>
<td>Safe Seat</td>
<td>-0.010 (0.003)</td>
<td>0.081 (0.017)</td>
<td>-0.045 (0.011)</td>
<td>0.016 (0.005)</td>
<td>0.002 (0.002)</td>
</tr>
<tr>
<td># Observations</td>
<td>[6060]</td>
<td>[6060]</td>
<td>[6060]</td>
<td>[6060]</td>
<td>[6060]</td>
</tr>
<tr>
<td>with Year Polynomial</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scandal × Safe Seat</td>
<td>-0.110 (0.037)</td>
<td>0.775 (0.241)</td>
<td>-0.090 (0.076)</td>
<td>0.117 (0.067)</td>
<td>-0.150 (0.048)</td>
</tr>
<tr>
<td>Scandal</td>
<td>-0.100 (0.022)</td>
<td>0.358 (0.116)</td>
<td>-0.182 (0.055)</td>
<td>0.156 (0.043)</td>
<td>-0.031 (0.022)</td>
</tr>
<tr>
<td>Safe Seat</td>
<td>-0.013 (0.003)</td>
<td>0.095 (0.017)</td>
<td>-0.054 (0.011)</td>
<td>0.015 (0.005)</td>
<td>0.001 (0.002)</td>
</tr>
<tr>
<td># Observations</td>
<td>[6060]</td>
<td>[6060]</td>
<td>[6060]</td>
<td>[6060]</td>
<td>[6060]</td>
</tr>
</tbody>
</table>

Each block presents the estimated coefficients on (i) the interaction between the dummy variable *Scandal* and the share of the presidential vote won by the incumbent’s party, *IPV*, and (ii) the dummy variable *Scandal*, from a separate regression. Robust standard errors in parentheses. Number of observations in brackets.