A Bayesian Explanation for Incumbency Advantage

Anthony Fowler
Harris School of Public Policy Studies
University of Chicago
anthony.fowler@uchicago.edu

Abstract

Incumbents appear to perform better in elections because they are incumbents, yet the most commonly proposed explanations for this phenomenon are unsatisfying and inconsistent with empirical evidence. I introduce a new explanation that is simple, parsimonious, and consistent with existing evidence. If voters lack perfect information about electoral candidates, incumbency is an informative signal of quality, and voters will update their beliefs accordingly—producing an incumbency advantage. I formalize these claims with a decision-theoretic model where voters receive noisy signals of candidate quality, and I experimentally test this explanation by providing voters with knowledge that removes the informative value of incumbency. When voters learn that their incumbent barely won office, they are less likely to support reelection and the incumbency advantage largely disappears, although the specific point estimates are imprecise. Furthermore, this theory, combined with historic time trends in voter knowledge, can explain the widely debated rise and fall of incumbency advantage at the end of the 20th century.

---

1 I thank Scott Ashworth, Chris Berry, Ethan Bueno de Mesquita, Mo Fiorina, Will Howell, Navin Kartik, Pablo Montagnes, Scott Tyson, Richard Van Weelden, Stephane Wolton, John Zaller, and seminar participants at APSA, UCLA, and UChicago for helpful comments and suggestions.
Incumbent electoral candidates in the U.S. and many other countries appear to perform better in elections simply because of their incumbency status. Scholars have long speculated about the sources of this incumbency advantage, but most of our current explanations are unsatisfying or inconsistent with empirical evidence. In this paper, I present a new explanation for incumbency advantage based on Bayesian learning under imperfect information. Typically, an incumbent candidate has won a previous election, leading voters to update their beliefs about her quality. If all available information, independent of incumbency, suggests that two candidates are of comparable quality, the voter will favor the incumbent, because the previous election provides an additional positive signal. This phenomenon, in the aggregate, will produce a positive incumbency advantage. I present a decision-theoretic model which formalizes these claims, and I empirically test its predictions with a survey experiment. Bayesian learning appears to explain a meaningful proportion of the observed incumbency advantage in American elections.

As a scientific matter, incumbency advantage is a prevalent feature of elections that we would like to better understand in order to learn more about democracy, political representation, and human behavior in general. Furthermore, our explanations for this phenomenon hold crucial implications for the health of democracy and for public policy. Some explanations implicate incumbents themselves who purportedly implement and exploit institutional advantages in order to keep themselves in power. For example, Fiorina (1977) argues that members of Congress maintain their incumbency advantage through case work, constituency service, and other institutional advantages, perpetuating a “system which subordinates the content of public policy” (p. 72). This notion is so powerful that it has been used to justify and advocate for term-limit laws which force incumbents to retire (e.g., Benjamin and Malbin 1992). Alternatively, an incumbency advantage driven by Bayesian learning is not such a cause for alarm. If anything, this explanation suggests that
incumbency advantage may be normatively desirable, arising because voters use incumbency as an informational signal in order to elect better leaders, on average.

First, I define incumbency advantage for the purposes of this study, and following previous studies, I estimate it across many offices and time periods. Next, I discuss previous explanations of incumbency advantage and explain why they are generally unsatisfying or inconsistent with empirical evidence. Because incumbency advantage is consistent across different offices in the U.S., its explanation is likely not specific to particular institutional features. Rather, a more general explanation based on the way that voters process information in all elections is more likely to explain this general phenomenon. Then, I present my own theory of incumbency advantage and formalize it with a decision-theoretic model. After explaining a key prediction of the model, I test it with a survey experiment. Informing voters of their incumbents' win margin virtually eliminates the estimated incumbency advantage. The results imply that most of the observed incumbency advantage in U.S. elections can be attributed to this particular mechanism, although the point estimates are imprecise. Finally, I examine historical trends in incumbency advantage and voter information, explaining how Bayesian learning can help us understand the dramatic rise and fall of incumbency advantage in the latter half of the 20th century.

Background on Incumbency Advantage

The term *incumbency advantage* has been used in the academic literature to describe several different phenomena and concepts. The term has been used casually to describe the fact that most incumbents who seek reelection succeed (e.g., Jacobson 2013). Of course, much of this phenomenon can be explained by partisan selection—Democratic incumbents tend to represent Democratic electorates and Republican incumbents tend to represent Republican electorates. *Incumbency advantage* can also be used to describe the extent to which incumbents outperform the
normal partisan voting behavior of their electorate (e.g., Ashworth and Bueno de Mesquita 2008). Part of this phenomenon can be explained by selection on candidate characteristics. Because they won a previous election, incumbents are probably better candidates than non-incumbents, even from the same party and running in the same place, leading them to outperform the normal vote their in their electorates. A third use of the term *incumbency advantage* refers to the extent to which incumbent candidates perform better in elections *because* they are incumbents. For the remainder of the paper, I will focus on this third use of the term *incumbency advantage*, because this is the particular phenomenon that currently lacks a compelling explanation. Of course, incumbents perform well in elections partly because they are good candidates who match the partisanship of the electorate, and previous work confirms that these forms of selection explain a meaningful portion of incumbent success (Alt, Bueno de Mesquita, and Rose 2011; Fowler 2014; Hirano and Snyder 2009; Samuelson 1987). However, we do not understand why a candidate performs better as an incumbent than as a non-incumbent running in the same place and time.

Varied uses of the term *incumbency advantage* create confusion in the empirical literature, as different research designs identify different quantities. For example, the retirement slump (Alford and Brady 1989; Ansolabehere and Snyder 2004) and panel regression estimates (Ansolabehere and Snyder 2002; Gelman and King 1990) fail to distinguish between the effects of incumbency and selection. These estimates are influenced by the extent to which new candidates may be lower quality than retiring incumbents. The sophomore surge (Erikson 1971; Levitt and Wolfram 1997) does attempt to exclude selection and identify the independent effect of incumbency, because the identity of the incumbent and non-incumbent candidate being compared is unchanged. However, other concerns such as time effects, mean reversion, and non-random retirement plague the interpretation of this approach. Regression-discontinuity (RD) designs (e.g., Lee 2008) can help to disentangle the effects of incumbency from selection by exploiting the quasi-random results of very close elections.
In expectation, the winners and losers of close elections should be comparable in terms of their characteristics, including their abilities and their partisan congruence with their constituency. However, estimating incumbency advantage using an RD design is less straightforward than it may initially seem (Erikson and Titiunik 2015; Fowler and Hall 2014). Nonetheless, evidence from RD designs suggests that incumbent candidates indeed perform better than they would in the counterfactual scenario where they ran in the same place and time but as a non-incumbent, although the mechanism behind this phenomenon is unclear.

To spark an initial discussion of potential mechanisms explaining incumbency advantage, Figure 1 presents RD estimates of incumbency advantage across different political offices in the U.S. and across different points in time in the 20th century. Specifically, these estimates arise from the fuzzy RD design employed by Fowler (2014). See the original paper for more details. The figure presents separate estimates for the U.S. House, the U.S. Senate, gubernatorial elections, statewide offices (lieutenant governors, secretaries of state, state attorney generals, etc.), mayors, and state legislators for all time periods for which data is available. The curves represent moving averages, where each point represents an estimate of incumbency advantage for all elections within 10 years (on either side) of a particular year.

[Figure 1]

First, we see that the incumbency advantage was positive at virtually every point in time in the 20th century and for every office for which we have available data. Incumbents appear to perform significantly better than they would in the counterfactual scenario where they were running as non-incumbents. Second, consistent with previous evidence (e.g., Ansolabehere and Snyder 2002; Gelman and Huang 2006; Jacobson 2015) we see that incumbency advantage rose dramatically between 1960 and 1990 and has since fallen. Third, consistent with Ansolabehere and Snyder (2002), we see that the magnitude of incumbency advantage and its changes over time are largely consistent
across offices. As pointed out by Ansolabehere and Snyder, the consistency of incumbency advantage across offices is particularly striking, because it casts doubt upon most of the previously proposed mechanisms or explanations for incumbency advantage.

**Previous Explanations for Incumbency Advantage**

The first studies of incumbency advantage and its rise over time focused on the U.S. House of Representatives (e.g., Erikson 1971; Mayhew 1974). Therefore, in trying to explain these phenomena, political scientists naturally thought about the unique features of the U.S. House and various changes that may have taken place since 1960. One salient feature of the U.S. House is that new districts are drawn every 10 years, and the redistricting process became much more politicized after Baker v. Carr and other Supreme Court decisions in the 1960s. Perhaps incumbents in the House increasingly used their political power to influence the drawing of their districts, creating a large and growing incumbency advantage (e.g., Cox and Katz 2002). Looking at Figure 1, estimates for the U.S. Senate cast doubt on the extent to which redistricting can explain incumbency advantage. While the estimates are less precise in the Senate, the level of incumbency advantage and the patterns of change over time almost exactly match those in the House, despite the fact that state lines are not regularly redrawn.

Perhaps the explanation has something to do with the institutional features present in both houses of Congress. Mayhew (1974) points out that members of Congress enjoy “franking privileges”—free mail to constituents, and the use of franked mail increased approximately four-fold between 1954 and 1970. Fiorina (1977) argues that members of Congress, starting in the 1960s, increasingly focused on case work and constituency service while shirking on policy. Further studies (e.g., McKelvey and Riezman 1992; Weingast and Marshall 1988) examine internal decision making in Congress and argue that members of Congress have set up various political institutions like
committees and a seniority system in order to ensure their reelection. However, Figure 1 casts doubt on the explanatory power of these factors as well. We see approximately the same level and rise of incumbency advantage for governors as we do for members of the House and Senate, despite the fact that governors do not conduct case work for constituents and enjoy none of these institutional advantages.

If incumbency advantage cannot be explained by institutional features of legislatures, perhaps it has something to do with the public salience of elected officials. After all, members of Congress and governors receive significant press coverage (e.g., Snyder and Stromberg 2010) and many voters are able to recall or recognize the names of these officials which may benefit incumbents when they run for reelection (e.g., Kam and Zechmeister 2013). Again, Figure 1 casts doubt on the importance of public salience, because we see similar levels and trends in incumbency advantage for statewide officials (secretaries of state, state directors of agriculture, etc.), mayors, and state legislators, all of whom receive less press coverage and enjoy less name recognition than governors and members of Congress. This evidence also casts doubt upon the view that the rise of television increased incumbency advantage through positive news coverage of incumbents (e.g., Prior 2006), because we see the same rise in incumbency advantage for low-salience elected officials who rarely appear on television. Further evidence suggests that greater television coverage of in-state elected officials has little effect on incumbency advantage (Ansolabehere, Snowberg, and Snyder 2006). This, however, does not preclude the possibility that television and changes in the media landscape influence incumbency advantage through less direct mechanisms, which I return to later in the paper.

Because incumbency advantage and its changes over time are so consistent across different offices in the U.S., any explanation that is particular to one or a few electoral settings is unlikely to explain a significant share of this phenomenon. Of course, the evidence in Figure 1 does not
preclude the possibility that incumbency advantage is driven by case work in the House, franked mail in the Senate, redistricting in state legislatures, television coverage for governors, local newspaper coverage for mayors, etc. However, Occam’s razor suggests that we should first consider more parsimonious explanations. Most likely, the primary explanations for incumbency advantage in U.S. elections are fundamental features of all of these electoral settings.

One prevalent feature of U.S. elections is campaign finance. Candidates need money to run their campaigns, although the amounts vary drastically across settings. Indeed, incumbents enjoy a significant financial advantage in elections (Fourrnaies and Hall 2014), and public funding laws which reduce this advantage also reduce incumbency advantage (Hall 2014). The point estimates from Hall (2014) suggest that the financial advantages of incumbency could explain approximately one-third of the electoral returns to incumbency. However, campaign finance alone is not a satisfying explanation for incumbency advantage. Why does incumbency produce more campaign contributions? Presumably, the financial advantage arises because access-oriented contributors believe there is an incumbency advantage independent of campaign finance that makes incumbents—even those who barely won last time—more likely to remain in office. By this logic, strategic campaign contributions might exacerbate an existing incumbency advantage (Ashworth 2006), but they cannot explain how the phenomenon originally arose.2

Another prevalent factor in virtually all elections is the strategic entry of candidates. The presence of an incumbent might “scare off” high quality challengers, who would prefer to wait until the incumbent retires before seeking office. Levitt and Wolfram (1997) estimate that the sophomore surge is smaller but still positive when the candidate who initially lost also runs again as a challenger.

2 This discussion assumes that access-oriented groups contribute to incumbents only when they are up for reelection, presumably so they can be in favor with the incumbent in their next term. However, they might also contribute at the beginning of the electoral cycle in the hopes of reaping rewards in the present term. This quid pro quo behavior, if widespread, could produce an incumbency advantage, but Hall’s (2014) estimates place an upper bound on the extent of this effect.
suggesting that challenger scare off is a source of incumbency advantage. However, one alternative explanation is that the losers only run again when the incumbent is expected to perform poorly. As a descriptive matter, experienced candidates are more likely to run in open-seat races than challenge a sitting incumbent, and experienced candidates tend to perform better than inexperienced candidates (Cox and Katz 1996; Jacobson 1989). However, utilizing a regression-discontinuity design, Hall and Snyder (2015) find that incumbency has no detectable effect on the experience level of the candidates in the next election. Perhaps experienced challengers avoid running against high quality incumbents, but they do not appear to be scared off by those incumbents who barely won their last election. At best, the evidence on the importance of challenger scare off is mixed. Levitt and Wolfram may overestimate its importance because of endogenous challenger entry, while Hall and Snyder may underestimate its importance by only focusing on a single, noisy measure of candidate quality. In any case, scare off does not appear to explain a large proportion of incumbency advantage. Furthermore, just like campaign finance, challenger scare off is not by itself a satisfying explanation. Presumably, high-quality challengers are scared off because of an incumbency advantage independent of challenger entry. Again, scare off might exacerbate an existing incumbency advantage but it could not independently produce the phenomenon.3

A simpler explanation for incumbency advantage that would apply broadly in many different electoral settings may arise from an important set of actors yet to be directly discussed: the voters.

3 One theoretically coherent explanation for incumbency advantage is that campaign contributors and potential challengers (and potentially other important players as well) play a game with multiple equilibria. In one equilibrium, high-quality challengers are scared off by incumbents, access-oriented contributors favor incumbents, and incumbency advantage is high. In another equilibrium, incumbency does not influence the behavior of potential challengers or contributors, and incumbency advantage is zero. In each case, no player has an incentive to deviate. This model can also conveniently explain why incumbency advantage varies significantly over time and across countries and might even be negative (different equilibria). I leave it to future work to flesh out such a model and assess it empirically. For now, I’ll point out that the estimates from Hall (2014) bound the explanatory power of this model, because estimates of the effect of public finance laws on incumbency advantage include the equilibrium responses from potential challengers and other players.
Many existing explanations for incumbency advantage explicitly or implicitly assume that voters are naive, simplistic, or downright foolish. For the explanations related to particularistic benefits, rational voters should not blindly support an incumbent who provides particularistic benefits if a newly elected representative would do the same. Rational citizens vote based on their expectations about the future, and they cannot credibly commit to rewarding the incumbent for effort in the previous term (Fearon 1999). For the salience explanations, the mechanism by which name recognition and media coverage should increase voter support is unclear. In the campaign finance account, voters are persuaded by campaign spending without realizing that, because of the financial incumbency advantage, incumbent spending should be less informative and less persuasive than challenger spending. However, incumbency advantage could also arise through the way in which rational voters process information and update beliefs when they lack the time or ability to perfectly observe the qualities of candidates. Furthermore, such an explanation could help us understand why incumbency advantage and its changes over time are so consistent across different offices in the U.S. Presumably, the informational challenges faced by voters and changes in their informational environment over time are similar regardless of whether they are evaluating candidates for Congress, governor, mayor, or the state director of agriculture.

A Simpler Explanation: Incumbency Is a Signal of Quality

In virtually every decision that we make, we lack perfect information about the choices. How should we make decisions in light of uncertainty? We simply aggregate all the information that we have and make our best guess. In cases where we have relatively little information, we may rely heavily on shortcuts or cues—noisy signals that contain information about the quality of the choices. For example, when deciding which movie to rent or which high-end restaurant to visit, we might rely upon discrete cues like the Academy Award or a Michelin star. For those who are serious movie
buffs or foodies, these cues will be relatively uninformative, because these decision makers are already well informed, but for the vast majority of consumers, an Academy Award or a Michelin star should lead them to positively update their beliefs about the quality of a movie or restaurant.

Elections are no exception. When rational voters decide between two candidates for elected office, they try to figure out which of the two will lead to better outcomes for them in the future (Fearon 1999). They often have little direct information about the quality of each candidate, so they must rely upon shortcuts and cues, some of which have been studied extensively. For example, voters might rely upon recent economic growth (Fiorina 1981), political endorsements (Lupia 1994), campaign spending (Gerber 1996), campaign activities and gaffes (Popkin 1991), sex scandals (Long 2011), and other crude indicators in order to learn about the quality of electoral candidates. One particularly informative signal about candidate quality is incumbency. If one of the candidates is an incumbent, then the voter knows that she won a previous election, which could be a highly informative signal about quality in the same way that Academy Awards and Michelin stars are informative signals about movies and restaurants. Each signal, while surely imperfect and error prone, informs the voter or consumer that many other people, at some point in the recent past, thought this candidate, movie, or restaurant was better than all the other options.

If voters know which candidate is an incumbent but don’t know the vote tally from the last election, then incumbency per se is an informative signal about the relative quality of the candidates. If all available information, independent of incumbency, suggests that two candidates are of comparable quality, then the voter will naturally favor the incumbent, because the previous election provides an additional positive signal. Even when available information suggests that the challenger is of higher quality, a voter might still support the incumbent in some circumstances. The extent to which these conflicted voters weigh the informative value of incumbency versus other signals will depend upon many factors including the strength of the other signals. In any case, this kind of
Bayesian information processing will produce an incumbency advantage where, all else equal, a candidate performs better in the counterfactual world where she is an incumbent. This phenomenon will even arise in situations where the incumbent barely won her last election—despite the fact that bare winners are, on average, no better than bare losers—because voters typically don’t know the exact vote margin from the last election. Voters can correctly conclude that, on average, incumbents are better than one would otherwise infer based on other information. Previous scholars have pointed out that informational cascades and social learning can help to explain fads, customs, and culture (Bikhchandani, Hirshleifer, and Welch 1992); herding in consumer purchasing (Banerjee 1992) and financial markets (Devenow and Welch 1996); and even polling effects and momentum in presidential primaries (Bartels 1988; Knight and Schiff 2010). In this paper, I argue that a similar process of social learning can explain incumbency advantage in elections.

A Decision-Theoretic Model of Elections with Noisy Signals about Candidate Quality

To formalize these ideas, I present a decision-theoretic model of elections with incumbents and noisy signals about candidate quality. As with all models, I do not intend to provide a complete description of elections. Rather, I present a theoretical abstraction which contains important elements of real-world elections and allows me to formalize and analyze a particular mechanism of incumbency advantage. After discussing this particular model, I will also discuss generalizations and the robustness of various theoretical results to different assumptions.

Consider a setting where candidate quality is binary—i.e., candidates can be either high or low quality (H or L). The proportion of high-quality types in the entire pool of candidates is π.

---

4 The model produces numerical results which are not intended represent actual numerical predictions in elections. Rather, the numerical results are useful only for identifying general phenomena and generating comparative static predictions.
where $0 < \pi < 1$.\textsuperscript{5} Voters receive independent, noisy signals (h or l) about each candidate which are wrong with probability $\varepsilon$, where $0 < \varepsilon < .5$.\textsuperscript{6} The number of voters is large enough that no individual vote is ever pivotal, but voters derive utility from supporting the candidate that they believe is more likely to be high quality. When indifferent, they independently flip fair coins.

Consider a sequence of two elections, an open-seat race and a second race where an incumbent seeks reelection. In the first election, two candidates are randomly drawn from the pool of candidates. Voters receive independent signals about each candidate (h or l) and cast their votes. Voters observe the outcome of the election but not the tally of votes. Then, in the next election, the initial winner seeks reelection as an incumbent. A challenger is drawn from the pool of candidates, independent of the initial winner and loser. Voters receive independent signals about the challenger (h or l) but no new signal about the incumbent. Lastly, voters cast their votes in the second election.

First, we can characterize voting behavior in the first election. As indicated above, voters are never pivotal and derive utility from supporting the candidate that they believe has the higher probability of being high quality. The voter can receive one of two possible signals for each candidate, h or l. If she receives a high signal for one candidate, she knows that one of two things occurred. Either the candidate is high quality and the signal was correct, which occurs with probability $\pi(1-\varepsilon)$, or the candidate is low quality and the signal was wrong, which occurs with probability $(1-\pi)\varepsilon$. Therefore, the probability that a candidate is high quality conditional on a high signal is

\begin{align*}
\frac{5}{\pi}
\end{align*}

\textsuperscript{5} $\pi$ is strictly bounded above 0 and below 1, because the cases where $\pi = 0$ or 1 are uninteresting. In these cases, all voters would be indifferent in both elections and the signals would be completely uninformative.

\textsuperscript{6} $\varepsilon$ is strictly bounded below .5 for two reasons. If $\varepsilon = .5$, then the signals are completely uninformative, and the model is uninteresting. All voters would be indifferent and flip coins in both elections. If $\varepsilon > .5$, then the signals will be negatively correlated with types, and receiving a high signal with $\varepsilon = \varepsilon^*$ would be identical to receiving a low signal with $\varepsilon = 1-\varepsilon^*$. In other words, considering values of $\varepsilon > .5$ is redundant, because the case where $\varepsilon = .6$ is identical to the case where $\varepsilon = .4$, $\varepsilon = .7$ is identical to .3, and so forth. Relatedly, the case where $\varepsilon = 0$ or 1 is uninteresting; the signals would no longer be “noisy” and there would be no uncertainty.
\[
\Pr(H|h) = \frac{\pi(1-\varepsilon)}{\pi(1-\varepsilon)+(1-\pi)\varepsilon}.
\]

This is a simple application of Bayes theorem. By the same logic, if a voter receives a low signal for a candidate, the posterior probability that the candidate is high quality is

\[
\Pr(H|l) = \frac{\pi\varepsilon}{\pi\varepsilon+(1-\pi)(1-\varepsilon)}.
\]

Because \(\Pr(H|h) > \Pr(H|l)\) for all possible combinations of \(\pi\) and \(\varepsilon\), voting decisions in the first election are straightforward. If a voter receives two high signals (hh) or two low signals (ll), she flips a fair coin. If a voter receives differing signals (hl), she will support the candidate for whom she received a high signal.

We can also characterize the aggregate vote shares in the first election. If both candidates are high quality (HH) or low quality (LL), then they will each receive half of the votes, in expectation, and one of them will be victorious by pure chance. For example, when both candidates are high quality (HH), \((1-\varepsilon)^2\) of the voters will receive the signals hh and flip coins. Similarly, \(\varepsilon^2\) of the voters will receive the signals ll and flip coins. \(\varepsilon(1-\varepsilon)\) of the voters will receive a high signal for the first candidate and a low signal for the second and support the first, while another \(\varepsilon(1-\varepsilon)\) of the voters will receive the opposite set of signals and support the second. Therefore, the expected vote share of each candidate will be \(.5(1-\varepsilon)^2 + .5\varepsilon^2 + \varepsilon(1-\varepsilon) = .5\). Alternatively, when the two candidates differ in quality (HL), \(2\varepsilon(1-\varepsilon)\) of the voters will receive the matching signals hh or ll and flip coins. \((1-\varepsilon)^2\) of the voters will receive the correct signals (hl) and support the high-quality candidate, while \(\varepsilon^2\) of the voters will receive two incorrect signals (lh) and vote for the low-quality candidate. As a result, the expected vote share for the high-quality candidate in this scenario is \(.5*2\varepsilon(1-\varepsilon) + (1-\varepsilon)^2 = 1-\varepsilon\). Therefore, the vote share, which voters do not observe, can take only two possible values. If both candidates are of the same quality, they each receive half the votes and one of them wins by chance.
If the candidates differ in quality, the high-quality candidate will win for sure with $1-\varepsilon$ of the vote share.

Voting behavior in the second election is more interesting, because the voter now has to consider both her private signals and the informative value of the previous election result. From the voter’s perspective, there are 8 possible sets of signals to consider. In the first election, she could have seen 4 different scenarios: hh—high signals for both candidates, ll—low signals for both candidates, hl—differing signals where the high-signal candidate won, and lh—differing signals where the low-signal candidate won. Then, in the second election, a new challenger appears for which the voter can receive either a high or low signal, producing 8 total sets of signals (hhh, hhl, llh, lll, hlh, hll, lhh, and lhl). For each set of signals, the voter calculates her posterior beliefs about the incumbent and the challenger and supports the candidate that she believes is more likely to be high quality. Posterior beliefs for the challenger are straightforward, because the calculation is the same as for new candidates in the open-seat race.

\[
\Pr(\text{challenger} = H| h) = \frac{\pi(1-\varepsilon)}{\pi(1-\varepsilon) + (1-\pi)^2\varepsilon} > \Pr(\text{challenger} = H| l) = \frac{\pi \varepsilon}{\pi \varepsilon + (1-\pi)(1-\varepsilon)}.
\]

Posterior beliefs about the incumbent are slightly more complicated, because in addition to the private signal that the voter received about the incumbent, she also knows that the incumbent is at least as good if not better than the losing candidate from the first election. As an example, consider the case where the voter received two high signals (hh) in the first election. There are three possible scenarios that could produce this set of signals. Both candidates could have been high quality (HH) and both signals were correct. This would occur with probability $\pi^2(1-\varepsilon)^2$. The incumbent could have been high quality while the losing candidate was low quality (HL) with the signal being correct only for the incumbent. This would occur with probability $2\pi(1-\pi)\varepsilon(1-\varepsilon)$. Lastly, both candidates could have been low quality (LL) and both signals were wrong. This only occurs with probability $(1-\pi)^2\varepsilon^2$. Note that the incumbent could not have been low quality while the
losing candidate was high quality, because she would not have won the election. Therefore, in two of the three possible scenarios, the incumbent candidate is high quality, and posterior beliefs about the incumbent are

\[
\Pr(Incumbent = H | hh) = \frac{\pi^2(1-\varepsilon)^2 + 2\pi(1-\pi)\varepsilon(1-\varepsilon)}{\pi^2(1-\varepsilon)^2 + 2\pi(1-\pi)\varepsilon(1-\varepsilon) + (1-\pi)^2 \varepsilon^2},
\]

which is greater than the posterior beliefs about the challenger for all possible values of \(\pi\) and \(\varepsilon\), even if the signal for the challenger is high. This result is easy to understand intuitively if we think about the information received by the voter. For both the incumbent and challenger, the voter received a high signal. However, the voter also knows that the incumbent defeated another candidate for whom she received a high signal, further raising her beliefs about the quality of the incumbent.

This intuition can help us understand the voter’s decision in each of the eight possible cases. In four of the eight cases, the voter receives the same signal for the incumbent and challenger. In these cases, the voter will always support the incumbent, for any possible values of \(\pi\) and \(\varepsilon\). Considering only the private signals, the voter is indifferent between the incumbent and the challenger. However, the previous election result leads the voter to increase her beliefs about the quality of the incumbent, and even a slight increase will lead the voter to strictly prefer the incumbent. In two of the eight cases, the voter receives a high signal for the incumbent and the low signal for the challenger. The voter now strictly prefers the incumbent for two reasons—the private signal is better for the incumbent and the previous election result leads the voter to further increase her posterior beliefs. In the final two cases, in which the voter receives a high signal for the challenger and a low signal for the incumbent, the voter’s decision is ambiguous. On one hand, the private signals suggest that the challenger is higher quality, in expectation, but the previous election result leads the voter to increase her beliefs about the incumbent. The relative magnitude of these
conflicting considerations will depend upon $\pi$ and $\varepsilon$, with some combinations of $\pi$ and $\varepsilon$ leading the voter to support the incumbent and others leading the voter to support the challenger.

Several theoretical results are presented graphically in Figure 2. The top-left panel shows how the optimal voting strategy varies across $\pi$—the proportion of high-quality candidates in the pool—and $\varepsilon$—the noisiness of the private signals. For any value of $\pi$, there is always a large enough value of $\varepsilon$ such that the voter will always support the incumbent regardless of her signals. This is because the informative value of the private signals is low enough to be swamped by the informative value of the of the election result. For any value of $\pi$, there is also an intermediate range of $\varepsilon$, such that the voter will support the incumbent for all possible sets of signals except llh, the worst case scenario for the incumbent. Lastly, if $\varepsilon$ is low enough, then the voter will support the incumbent for all sets of signals except llh and lhh, cases where they received a low signal for the incumbent and a high signal for the challenger.

[Figure 2]

Having characterized voting behavior in the second election, we can calculate incumbency advantage in this theoretical setting. An empirical researcher would say that there is an incumbency advantage in this setting if incumbents who won their first election in a tie receive more than half the votes, on average, in the second election when they seek reelection. One nice feature of this simplified model with binary candidate quality is that there are tied elections—where the expected vote share of each candidate is .5 and one of the candidates wins by the luck of the aggregations of coin tosses. In fact, ties occur with probability $\pi^2 + (1-\pi)^2$. Therefore, we can mimic the RD design and calculate the expected vote share of the incumbent conditional on a previous tie for all possible values of $\pi$ and $\varepsilon$. For the darkest region of values in Figure 2 where voters always support the incumbent, the incumbent vote share conditional on a tie is obviously 1. For the next region, this vote share will be 1 minus the probability that a voter receives the signals llh, conditional on a tie.
For the lightest region, the vote share is 1 minus the probability that a voter receives the signals llh or lhh, conditional on a tie.

The top-right panel of Figure 2 presents this quantity across all possible values of \( \pi \) and \( \varepsilon \). The vote share for incumbents following a tie is greater than 0.5 for all possible values of \( \pi \) and \( \varepsilon \). In other words, the incumbency advantage is always positive in this model. Generally speaking, incumbency advantage is greater for extreme values of \( \pi \) where there is an uneven mix of good and bad candidates in the pool and for large values of \( \varepsilon \) where private signals are less informative than incumbency. Even in a setting where there is no institutional benefit to incumbency, no learning from experience, no scare off of high-quality challengers, no campaign finance, etc., incumbency advantage can arise as voters use a prior election result to update their beliefs about the quality of the incumbent.

Incumbency advantage arises in this model, because the previous election result is a meaningful signal about the quality of the incumbent. I assume that voters know who the incumbent is, i.e., they know the previous election result, but they do not know the vote share from the previous election. The previous vote share of the incumbent could have been 1−\( \varepsilon \), in which case the incumbent is high quality for certain, or the previous vote share could have been 0.5, in which case the incumbent’s quality could be high or low. When the previous election was tied, the voter is unaware of this fact, and she gives some weight to the possibility that the incumbent won with 1−\( \varepsilon \) of the vote share, meaning that even those incumbents who were no better than their initial opponent benefit from this Bayesian information processing.

This insight provides a method by which this particular explanation for incumbency advantage can be tested empirically. If voters are unaware of the previous vote share and if Bayesian information processing produces an incumbency advantage, this advantage should be reduced if voters become aware of the incumbent’s vote share in her previous election. In the model above, the
voter knows her incumbent’s quality was either equal to or greater than the quality of her initial opponent, but if the voter becomes aware that the last election was a tie, then she knows that her quality was exactly equal to the previous opponent, which should decrease her posterior beliefs about the chances that the incumbent is high quality. To see this, reconsider the case where the voter receives two high signals in the first election for which the voter’s posterior beliefs were written above. If voters do not observe the vote share, there are three possible scenarios that could have produced this set of signals, two of which mean that the incumbent is high quality. However, if the voter learns that the initial election was a tie, one of those positive scenarios (high-quality incumbent defeated a low-quality opponent) is ruled out, reducing the voter’s posterior beliefs:

\[
\Pr(Incumbent = H|hh, tie) = \frac{\pi^2 (1-\varepsilon)^2}{\pi^2 (1-\varepsilon)^2 + (1-\pi)^2 \varepsilon^2} < \Pr(Incumbent = H|hh).
\]

To what extent should incumbency advantage be reduced if voters become informed about the previous vote share? To address this question, I analyze an alternate version of the model above where voters observe the vote share in the first election. Behavior in the first election is unchanged, because voters have the same information and they simply support whichever candidate they believe is more likely to be high quality, flipping coins if they receive the same signal for both candidates. Then, however, voter behavior in the second election is quite different, because voters can observe whether their incumbent received only half the vote in the initial election or \(1-\varepsilon\). If the incumbent received a vote share of \(1-\varepsilon\), the voters knows that she is high quality and reelects her for certain, because she can never be certain of the challenger’s quality. However, the more interesting case for studying incumbency advantage occurs when the voter observes a virtual tie in the initial election. The middle-left panel of Figure 2 characterizes the voter’s optimal behavior in the second election after observing a tie in the first election. Unlike the previous model where vote share was unobserved, there is no set of signals for which the voter would always support the incumbent,
regardless of $\pi$ and $\varepsilon$. Specifically, there are 5 regions of $\pi$ and $\varepsilon$ for which the optimal voting strategy differs.

When $\pi > .5$, meaning that most candidates are high quality, and $\varepsilon$ is large enough, the voter will reelect the incumbent regardless of her signals. Alternatively, when $\pi < .5$ and $\varepsilon$ is large enough, the voter will support the challenger regardless of her signals. This can be understood by considering situations where signals are very noisy ($\varepsilon$ just less than .5). As $\varepsilon$ approaches .5, conditional upon a tie, the probability the incumbent is high quality converges to $\frac{\pi^2}{\pi^2 + (1-\pi)^2}$ which is greater than $\pi$ when $\pi > .5$ and less than $\pi$ when $\pi < .5$. The tie signals that the winner and loser of the initial election had the same level of quality. When most candidates are high quality, a tie between two high-quality candidates is particularly likely, and when most candidates are low quality, a tie between two low quality candidates is particularly likely. Eggers (2015) discusses a similar theoretical result whereby an asymmetric distribution of candidate quality can produce a positive or negative incumbency advantage.

For intermediate values of $\pi$ or for low values of $\varepsilon$, the previous election result contains little information relative to the voters’ signals. In these cases, the voter will support the incumbent if she receives one of the four sets of signals that favor the incumbent (hhh, hhl, hll, and lhl), and she will support the challenger if she receives one of the four sets of signals that favor the challenger (hlh, lhh, llh, lll). For high values of $\pi$ and intermediate values of $\varepsilon$, the voter will support the incumbent for all signals except llh, and for low values of $\pi$ and intermediate values of $\varepsilon$, the voter will support the challenger for all signals except hhl.

Having characterized voting behavior in the second election when the voters observed a tie in the first election, we can characterize incumbency advantage in this particular model. The middle-right panel of Figure 2 presents the expected incumbent vote share following an initial tie for all possible values of $\pi$ and $\varepsilon$. Unlike the previous model, there is no systematically positive incumbency
advantage. If $\pi > .5$, incumbency advantage is positive, and if $\pi < .5$, incumbency advantage is negative. In general, incumbency advantage increases as $\pi$ increases, and the absolute value of incumbency advantage increases as $\varepsilon$ increases, but the latter rule is not universal.

For the purposes of generating empirical predictions, the most relevant quantity is the difference in incumbency advantage when voters do or do not observe the previous vote share—the differences between incumbency advantages observed in the top-right and middle-right panels, shown graphically in the bottom-left panel of Figure 2. The effect of informing voters about the previous vote share is weakly negative for all possible values of $\pi$ and $\varepsilon$. In other words, if incumbency advantage is driven by this kind of Bayesian information processing, this phenomenon will be reduced when voters are informed of their incumbents’ previous vote share, a prediction that can be tested experimentally.

An incumbency advantage driven by Bayesian updating generalizes beyond the particular structure of this model. I have formally examined a model where candidate quality and signals are binary. However, at its core, this model relies on much more general logic. The vote share from a previous election is a noisy measure of the relative quality of the candidates, representing an aggregation of signals across the electorate. If a voter knows that a candidate’s vote share in a previous election was above some threshold, she should raise her posterior beliefs about the quality of the candidate, integrating over all possible values above that threshold. This is the rationale for incumbency advantage. If a voter further learns that a candidate’s vote share was exactly at that same threshold, she should now lower her posterior beliefs, but the extent of this decrease could be greater or less than the initial increase. This is the rationale by which informing voters about the vote share should reduce incumbency advantage. These predictions apply weakly for all possible distributions of candidate quality or voter signals.
To summarize the theoretical results, incumbency advantage can arise from an abstract electoral setting where voters process information under imperfect information. Even when there are no institutional advantages to incumbency, returns to experience, or other factors that previous scholars have discussed, voters will use the previous election result to update their beliefs about the quality of the incumbent, leading bare winners to perform better when they seek reelection. Furthermore, if this particular mechanism is relevant for American elections, incumbency advantage should be reduced when voters learn the previous vote share of their incumbent.

Experimental Evidence

The primary empirical prediction from the model above is that when voters observe the vote share from the previous election, the estimated incumbency advantage decreases. To test this prediction, I embedded a survey experiment in a module of the 2014 Cooperative Congressional Election Study (CCES). The module surveyed a nationally representative sample of 1,474 eligible voters in the weeks leading up to the midterm elections of November 2014. All surveys were conducted online by YouGov. My experiment focused specifically on the 2014 gubernatorial election, because gubernatorial races are typically more competitive than Congressional races, and high proportion of respondents resided in a state that had last elected its current governor by a close margin. I exclude from my analysis the 14 states that did not hold a gubernatorial race in 2014. I also exclude the four states in which the incumbent governor in 2014 was term limited and therefore ineligible to seek reelection (Arizona, Arkansas, Maryland, and Nebraska). Unfortunately, I also exclude Oklahoma, because YouGov failed to ask the relevant questions in this state. This leaves 1,049 eligible voters in 31 states for my analysis. For my initial analysis, I do not exclude the three states where the incumbent governor was eligible to run but did not (Massachusetts, Rhode Island,
and Texas), because these states are necessary to estimate incumbency advantage if retirement is strategic.

In designing the experiment, I attempted to test my theoretical prediction in the simplest possible way. A randomly selected half of the respondents—the control group—received no information about their governor. The other half of respondents—the treatment group—saw the following text on a separate page in the middle of their survey: “We’d like to provide you with information on the last gubernatorial election in your state. In the last election, your governor received $X\%$ of the two-party vote.” $X$ was replaced by the actual two-party vote percentage for their state. A previous task had clarified the meaning of the two-party vote, so respondents should have known that the lowest possible value is 50 percent—meaning that their current governor won in a virtual tie, and the highest possible value is 100 percent. After completing approximately 10 unrelated survey questions and reading tasks, respondents were asked which gubernatorial candidate in the upcoming race they preferred: “In the race for Governor in your state, who do you prefer?” Respondents could pick from the names of the major candidates, and they could also select “Other”, “I’m not sure”, or “No one.” Using this survey question, I first code a variable which takes a value of 1 if the respondent supports the Democratic candidate and 0 if the respondent supports the Republican candidate. For those respondents that did not support either candidate, I code them as 0.5 so as to avoid dropping observations, which is particularly important if the treatment influenced whether the respondent supported one of the major candidates.

This simple experiment and this single vote choice question are sufficient to test the primary prediction of the model. I would like to estimate incumbency advantage for the control and treatment groups and compare these estimates. To do this, I utilize the same approach for estimating incumbency advantage as employed in Figure 1. Specifically, I employ a fuzzy RD design, incorporating the actual election returns from the most recent gubernatorial election in each
respondent’s state as the running variable. Unfortunately, this approach is statistically imprecise, but it offers the most direct test of the theoretical prediction. For the control group, I estimate an incumbency advantage of 4.9 percentage points (standard error = 3.7)—strikingly close to the recent incumbency advantage estimates from Figure 1 generated from actual election returns. For the treatment group, this estimate shrinks to 0.5 percentage points (standard error = 3.5). Therefore, the estimated incumbency advantage is 4.4 percentage points lower (standard error = 5.1) for the treated respondents than for the control group. Although the point estimates are imprecise, they suggest that almost all of the incumbency advantage observed in elections disappears when voters become informed of their incumbent’s vote share from the previous election.

Figure 3 illustrates this effect graphically. The vertical axis represents support for the Democratic candidate as measured in the survey. The horizontal axis plots the two-party vote share received by the Democratic candidate in the most recent election, as measured by official election returns. Each data point shows the average level of Democratic support for each state and each treatment condition. The curves represent fourth-order polynomial fits of the data, weighted by the number of respondents in each state. For the control group, we see a large, discontinuous jump in support for the Democratic candidate as we move from states that barely elected a Republican in the last election to those that barely elected a Democrat, indicating a large, positive incumbency advantage. For the treatment condition, we see no such discontinuity, suggesting that the experimental treatment decreases incumbency advantage.

Figure 4 presents another, perhaps simpler way to assess the effect of the treatment on incumbent support. I drop the states where no incumbent is on the ballot, and I recode the vote choice variable to indicate support for the incumbent candidate, again leaving those not supporting a major candidate in the middle of the scale at 0.5. To assess the effect of the treatment on incumbent
support, I regress this variable on a binary variable indicating treatment status, and to improve precision, I also include state fixed effects and respondents’ pre-treatment support for the incumbent, which was measured in the common content of the survey, well before the experimental module. However, estimating the average effect of the treatment on incumbent support is uninteresting. My theory predicts that informing voters about the previous vote share will reduce incumbency advantage for those incumbents that barely won their previous election in a virtual tie, but the effect of this information for all incumbents is theoretically ambiguous. Therefore, Figure 4 presents the average effect of the treatment for different levels of previous win margin, with the prediction that this effect should be negative for the smallest win margins and should increase as the window of included observations is widened. Indeed, the treatment appears to decrease incumbent support by more than 4 percentage points for those states where the incumbent previously won with less than 51 percent of the two-party vote, and this effect gradually dissipates as the win margin is increased.

[Figure 4]

Although the results of the experiment are not statistically precise, the estimated effects are always in the direction predicted by the theory that incumbency advantage is driven by Bayesian information processing under imperfect information. The point estimates suggest that this model could explain a large proportion of the observed incumbency advantage, because simply informing voters of the previous vote share nearly eradicates the incumbency advantage altogether. Furthermore, no previous explanations or theories of incumbency advantage would predict that such a simple, informational treatment would have this effect. I plan to conduct more precise tests of these hypotheses in a future, higher-power experiment, but I interpret these initial results as strong, suggestive evidence consistent with my simple, informational theory.
Explaining Variation in Incumbency Advantage over Time

If Bayesian information processing contributes to a large incumbency advantage today, what can explain the dramatic rise and fall of incumbency advantage in the latter half of the 20th century? As previously discussed, the explanation is unlikely to stem from particular institutional reforms, because the trends in incumbency advantage are so consistent across different offices which have not been subject to the same reforms. However, if incumbency advantage is largely explained by information processing under uncertainty, then we could easily explain a consistent shift in incumbency advantage. Any change in the overall level of voter information or the way that voters process that information would cause incumbency advantage to shift in a similar way across all offices.

In particular, the overall level of political information may be particularly important for predicting the extent of incumbency advantage. If voters are already highly informed about the candidates without considering incumbency, then the previous election result will have little additional informative value, and incumbency advantage will be low. Alternatively, if voters are generally uninformed about political candidates, the previous election result will be an informative signal of candidate quality, and incumbency advantage will be high. Therefore, the rise and fall of incumbency advantage could be simply explained by Bayesian reasoning and a general decrease and then increase in voters’ knowledge about political candidates.

To my knowledge, the best available measure of voters’ knowledge of political candidates over time comes from the American National Election Studies (ANES). In 1958 and in every even year from 1964 to 2000, the ANES asked a nationally representative sample of eligible voters to recall the name of any candidate running for the U.S. House of Representatives in each respondent’s district, and the surveyor recorded whether the respondent was able to correctly recall at least one of the candidates. Figure 5 plots the average recall rate in each year alongside the average estimate of
incumbency advantage from Figure 1, pooling across all offices. Consistent with the theoretical prediction, the dramatic rise in incumbency advantage from 1960 to 1980 was accompanied by an equally dramatic fall in voter knowledge of political candidates, and the drop in incumbency advantage since the mid-1990s was accompanied by a rise on voter knowledge. Of course, the estimates of candidate recall in any given year can be imprecise, but the general trends over time almost exactly mirror, in the opposite direction, the trends in incumbency advantage.

[Figure 5]

The combination of this evidence and theory suggests that the rise and fall of incumbency advantage could be explained by widespread changes in the level of information that voters had about political candidates. Before 1960, voters were highly informed, and they had little use for the crude signal of incumbency. However, as voters became less informed, for whatever reason, they increasingly relied upon the previous election result to draw inferences about candidate quality. In recent years, again for whatever reason, voters became more knowledgeable about political candidates, but still not as knowledgeable as they were around 1960, and the incumbency advantage accordingly fell.

The causes of these changes in voter knowledge are beyond the scope of this particular paper but worth some discussion nonetheless. What can explain the fall and rise of voter knowledge over this time period? Changes in the media landscape offer one likely set of explanations. The introduction of television caused many voters to substitute away from information-rich newspapers, thereby decreasing the voter information and participation (Gentzkow 2006) and leading voters to focus more on national vs. local races (Song 2014). Subsequent changes in the media landscape such as cable televisions and the internet had important political effects (e.g., Prior 2005), which may have, on net, increased voter knowledge about candidates, especially among active voters.
This account presents an interesting paradox about television and incumbency advantage. Television coverage cannot directly explain incumbency advantage or its rise, because we see the same trends in incumbency advantage for offices that receive little television coverage (Ansolabehere and Snyder 2002), and within a particular race, incumbents are just as popular in the parts of their electorate receiving out-of-state television coverage (Ansolabehere, Snowberg, and Snyder 2006). However, if television generally decreases voter knowledge of candidates, as previous evidence suggests, and if decreased knowledge increases incumbency advantage as Bayesian learning would predict, then television could have indirectly caused much of the rise in incumbency advantage after 1960. Similarly, other changes in the media landscape that later increased voter knowledge may have caused the subsequent fall of incumbency advantage. This insight may even help to resolve some of the apparent disagreements in the literature. Even if the actual television coverage of candidates has little effect on incumbency advantage (e.g., Ansolabehere, Snowberg, and Snyder 2006), more television viewership in the electorate may nonetheless be good for incumbents (e.g., Prior 2006), because television indirectly leads voters to be less informed about the candidates and therefore rely more heavily upon incumbency in forming beliefs about candidate quality.

Discussion and Conclusion

Incumbency advantage is a widely studied and substantively important phenomenon in electoral politics. Despite decades of research, most of our current explanations for this phenomenon are theoretically unsatisfying or inconsistent with empirical evidence. This paper offers a new, parsimonious explanation based on the way in which rational voters process information in light of imperfect information. An incumbency advantage can arise when imperfectly informed voters use the previous election result to update their beliefs about candidate quality. A directly testable implication of this theory is that incumbency advantage would decrease if voters became
aware of the precise vote share in the previous election, removing the informative value of incumbency *per se*. A survey experiment in the run-up to the 2014 gubernatorial elections confirms this prediction. When a random subset of voters become informed of their governor’s vote share in the most recent election, support for incumbents who were barely elected decreases and the estimated incumbency advantage virtually disappears. This theory can also help to explain the dramatic rise and fall of incumbency advantage in the latter half of the 20th century. This phenomenon was accompanied by a sharp drop and then rise in voter’s knowledge of political candidates, and the theory predicts that incumbency advantage will be greater when voters are less knowledgeable as the relative informative value of incumbency increases.

The Bayesian explanation for incumbency advantage offered in this paper may also be useful in explaining the significant variation in incumbency advantages across different countries. In general, incumbency advantages appear to be large in well-developed countries like the U.S., the U.K. (Eggers and Spiirling 2014), Germany (Hainmueller and Kern 2008), Australia (Horiuchi and Leigh 2009), and Canada (Kendall and Rekkas 2012), and incumbency advantages appear smaller in less-developed settings like India (Uppal 2009), Brazil (Klasnja and Titiunik 2015), Zambia (Macdonald 2014), and the U.S. and U.K. in the 19th century. Bayesian learning offers a simple explanation for this pattern. In the model, incumbency advantage arises, because the previous election result informs a voter about the private signals of other voters. However, the informative value of an election result will be diminished in the presence of electoral fraud, where voters can be less confident that the election results reflects the true preferences of other voters rather than the ability of the winning candidate to manipulate elections. Because electoral fraud and manipulation are more common in less developed countries (Lehoucq 2003; Simpser 2013), voters in these settings may reasonably rely less heavily upon incumbency—or even assign negative weight to incumbency, thereby producing a small or even negative incumbency advantage in these settings.
When elections are largely fair, we would expect a positive incumbency advantage from Bayesian learning, but if elections are marred by fraud and manipulation, rational voters should not positively update their beliefs about the quality of incumbents, and the incumbency advantage will be smaller.

This model of an incumbency advantage driven by rational information processing has important implications for our normative evaluations of democratic elections. Many scholars, explicitly or implicitly, interpret incumbency advantage as a sign of a broken political system that incumbents exploit in order to keep themselves in power, and policy advocates cite incumbency advantage as a justification for term limit laws. However, if the explanation presented here is correct, incumbency advantage is a byproduct of optimal voter behavior. When voters’ private signals are noisy and imperfect, they should use incumbency to update their beliefs. In doing so, imperfectly informed voters can, on average, select higher quality leaders. Sometimes, low-quality incumbents benefit from this behavior, thus producing an incumbency advantage, but nonetheless, voters are better off when they incorporate incumbency into their voting calculus. Of course, voters would be even better off if they had perfect information about candidate quality to begin with, but such a goal is impractical and unrealistic. Conditional upon being imperfectly informed, incumbency is a valuable signal of quality.

The theory and evidence presented in this paper suggests that the explanation of incumbency advantage, the subject of much debate and theorizing, may turn out to be embarrassingly simple. If voters aim to select the best candidate, and if they have little information available, incumbency represents a valuable signal of quality that can and should influence vote choices. The incumbent won at least one previous election, so all else equal, her expected quality is greater than an untested candidate. As a result of Bayesian information processing of rational but imperfectly informed voters, incumbent candidates can expect to perform better in elections simply because they are incumbents.
References


Hall, Andrew B. 2014. How the Public Funding of Elections Increases Candidate Polarization. SPSA 2014 Annual Meeting Paper.

Hall, Andrew B. and James M. Snyder, Jr. 2015. How Much of the Incumbency Advantage is Due to Scare-Off? *Political Science Research and Methods*, forthcoming.


The figure presents average estimates of incumbency advantage for different offices (or groups of offices) over time. All estimates of incumbency advantage arise from the fuzzy RD design of Fowler (2014), exploiting close elections and accounting for which incumbents seek reelection. Each data point represents an average estimate including the ten years before and after that particular year. The incumbency advantage and its changes over time appear to be quite consistent across different offices in the U.S., casting doubt upon most previous explanations of incumbency advantage.
The pictures above present several different results from the decision-theoretic model. The top-left panel shows, for all possible values of π and ε, which sets of signals will lead a voter to support the incumbent. The top-right panel shows, for all possible values of π and ε, the expected vote share of incumbent candidates who won their previous election in a virtual tie—analogous to RD estimates of incumbency advantage. Because this value is well over .5 for all possible values of π and ε, there is a positive incumbency advantage in this model. The middle-left and middle-right panels mimic the top-left and top-right panels, respectively, but for an alternative version of the model where voters observe the vote share in the initial election. The bottom-left panel shows the weakly negative effect of informing voters about the vote share on incumbency advantage. To summarize, Bayesian learning under imperfect information produces a positive incumbency advantage which decreases if voters are aware of the precise vote tally from the previous election.
Figure 3. Informing Voters about Election Results Decreases Incumbency Advantage

The graph shows that the estimated incumbency advantage is large for the control group but virtually disappears for the treatment group, which was informed of their incumbent governor’s vote share in the most recent election. The vertical axis indicates support for the Democratic candidate in the CCES survey. The horizontal axis represents the two-party vote share received by the Democratic candidate in the last gubernatorial election in each respondent’s state. The points represent average values for all individuals within a particular state and treatment group, the curves indicate separate fourth-order polynomial fits (weighted by the number of respondents in each state) for the treatment and control groups, and the dotted lines indicate standard errors. The discontinuity at the electoral threshold is large for the control group, indicating a large, positive incumbency advantage, and this discontinuity virtually disappears for the treatment group.
Figure 4. Effect of Treatment for Different Degrees of Electoral Closeness

The figure shows estimates of the average effect of the treatment on incumbent support for all states where the previous incumbent vote share was less than 50, 51, 52, etc. percentage points. Specifically, each data point results from a regression of incumbent support (which takes possible values of 1, .5, and 0) on a treatment indicator, pre-treatment incumbent support, and state fixed effects. Dotted lines indicate standard errors. The treatment appears to decrease incumbent support by over 4 percentage points for those states where the previous incumbent win margin was less than one point, and as expected, this effect gradually dissipates as the previous win margin increases.
Figure 5. Incumbency Advantage and Voter Information over Time

The gray curve plots an average estimate of incumbency advantage over time using the same empirical method as in Figure 1 but pooling across all offices. The black curve plots a proxy for voter knowledge over time—the proportion of respondents in each ANES survey that can correctly recall the name of at least one candidate in the upcoming election for U.S. House. Incumbency advantage rose and then fell and the same time that voter information dropped and then rose. This is consistent with the idea that incumbency advantage is driven by Bayesian information processing. When voters are highly informed, they already have strong signals of candidate quality, and the relative value of the previous election result as an informative signal is low. However, when voter information is low, the informative value of incumbency is high, and voters will increasingly rely upon incumbency as a signal of quality. These observed trends in voter information over time combined with Bayesian learning can explain the rise and fall of incumbency advantage over time.