

Electoral Selection and the Incumbency Advantage*

Scott Ashworth[†]

Ethan Bueno de Mesquita[‡]

First Version: August 13, 2004

This Version: August 13, 2004

1 Introduction

Sitting members of Congress exhibit an extraordinary degree of success when they seek reelection. Explaining the source of this incumbency advantage has been one of the most active topics in the study of American politics over the past quarter century. The dominant explanation is that the incumbency advantage is caused by the ability of members of congress to provide constituency service (e.g., Fiorina (1977) or Cain, Ferejohn, and Fiorina (1987)). Members of Congress, the argument goes, provide local public goods to their constituents and thereby shore up support for themselves. Indeed, Levitt and Snyder (1997) and Fiorina and Rivers (1989) find evidence that incumbents who provide more pork-barrel spending or have greater district presences are more successful in their reelection contests. Recently, Cox and Katz (2002) suggested that redistricting also plays an important role in explaining the electoral success of incumbents.

While these explanations are surely an important part of the puzzle, Ansolabehere and Snyder (2002) have reported a set of new empirical findings that show they cannot be the full story. They demonstrate that the trend displayed by the incumbency advantage for members of the U.S. House of Representatives was actually shared by *all* offices elected in state-wide

*We are indebted to Andrew Martin for teaching us how to do simulations. We also benefited from the comments of Howard Rosenthal, Steve Smith, and the seminar audience at the Kellogg School of Management.

[†]Department of Politics, Princeton University. Email: sashwort@princeton.edu

[‡]Department of Political Science, Washington University. Email: ebuenode@artsci.wustl.edu

elections. In particular, the advantage was small during the 1940s and 50s, grew sharply during the 1960s, and stabilized at a high level during the 1970s, 80s, and 90s. Both the constituency service and the redistricting arguments are insufficient to explain these findings. Other than Congress-people, none of the offices that Ansolabehere and Snyder studies experienced redistricting. Indeed, since most of them were state-wide offices (Senators, Governors, etc.), they could not be re-districted. Further, many of the offices that Ansolabehere studied were positions that involved no opportunity to provide constituency service (e.g., state auditors). As such, Ansolabehere and Snyder's empirical work indicates the need to explore additional explanations of the incumbency advantage.

The literature discusses at least two explanations of the incumbency advantage that are capable of explaining the evidence in Ansolabehere and Snyder (2002). The electoral selection hypothesis is based on the claim that voters elect candidates based on an assessment of some characteristic of the candidate that the voters' find attractive. As such, the fact that a candidate has been elected in the past means that the candidate possess some trait that impressed the voters in the past. Such a "high ability" candidate should be expected to do well in the future. Another explanation of the incumbency advantage is that challenger quality is greater in open seat races than in races with incumbents. Levitt and Wolfram (1997), Ansolabehere and Snyder (2002), and Gowrisankaran, Mitchell, and Moro (2003) all find that both selection and challenger quality are important, accounting for about 1/2 of the advantage each. In this paper, we focus on selection.

We study the comparative static properties of an electoral selection model. We show that if voters get more precise information about the incumbent's ability, then the incumbency advantage is greater, while more variability in the ideological stance of a district reduces the incumbency advantage. We also show that if a district becomes more competitive, in the sense of being more balanced between the two parties in open seat elections, then the incumbency advantage increases.

Here is a quick preview of the analysis. We build an electoral selection model in which both legislative institutions and the ideological distribution of the electorate across districts can vary. The incumbency advantage arises in the model because voters get information about the incumbent, and use that information to select high ability candidates. More formally, since

voters vote for candidates based on merit, in each election the left tail of the distribution of candidate “ability” is censored. This implies that candidates who have won elections before will have greater expected abilities than either challengers or candidates in open seat elections. This point is made formally in Ashworth (2001), and is shown to provide a good fit to Congressional elections in simulations by Zaller (1998) and in a structural econometric model by Gowrisankaran et al. (2003).¹

The comparative statics flow from changes in the quality of information or from the willingness of voters to act on information. With a more precise signal, the voter is better able to identify high-ability incumbents. This enhances the censoring effect that leads to the advantage in subsequent elections. On the other hand, more variable shocks to the voter’s ideal point or more district-level bias toward one party both make it more likely that a voter will face a trade-off between choosing a high-ability candidate and choosing an ideologically sympathetic candidate. Since this trade-off is sometimes resolved in favor of ideology over ability, the censoring is less extreme, and the incumbency advantage is reduced.

We show that these comparative static results allow the selection model to explain both Ansolabehere and Snyder’s historical findings and the central comparative facts about incumbency advantages. This is a crucial theoretical test for the selection model, since the great strength of the traditional constituency service model is the way it provides a unified account of a variety of facts, from both comparative and American politics. In particular, empirical research has shown that the incumbency advantage is greater in (majoritarian) presidential systems than in (majoritarian) parliamentary systems (compare Gelman and King (1990) and Katz and King (1999)). The constituency service theory explains these patterns by pointing out that MPs have fewer resources, and do less constituency service, than do MCs (Cain et al., 1987). Similarly, the constituency service model explains the rise in the incumbency advantage through the increase in constituency service.

In this paper, we show that the electoral selection model can also account for these comparative and historical facts, as well as others less easily explained by constituency service, such as the increase in non-legislative incumbency advantages. We also look at additional institutional factors: Cox and Morgenstern (1995) and Katz (1986) find that the incumbency

¹Of course, the basic idea is older. Zaller (1998) attributes it to a seminar by Doug Rivers as far back as 1988.

advantage is greater in states with single-member districts than in those with multi-member districts, and Alesina, Fiorina, and Rosenthal (1991) report that the incumbency advantage is stronger in states that send divided delegations rather than unified delegations to the Senate. We show that the selection view can account for *all* of these facts. While the model is stylized and incomplete, the fact that so many of the comparative statics are right should increase our confidence that electoral selection of the sort we model is one of the factors contributing to the incumbency advantage.

The basic logic is as follows. Compared to a voter in a multi-member district, a voter in a single-member district has more information, since she does not have to sort out the contributions of several representatives to any good signal she observes. Compared to a voter in a parliamentary system, a voter in a presidential system can make better use of her information, since her vote for the legislature is not bundled with her vote for the executive. In both cases, the single-member district, presidential system's voter has more effective information and, consequently, the censoring of the ability distribution is more extreme. This leads to a greater incumbency advantage.

The comparison of presidential and parliamentary incumbency advantages is driven by differences in the willingness of the voter to replace low ability incumbents. A similar mechanism explains changes over time in the U.S. The period since the mid-60s has been one of greater competitiveness between the parties in most states. The Democratic deadlock on the South was broken, and Republican strength declined in the North. For House elections, the decline of the Northern Republican party was exacerbated by redistricting in the wake of *Baker v. Carr* (Cox and Katz, 2002). These changes made the pivotal voter in most state more willing to swing from one party to the other, which implies more responsiveness to information about incumbent quality and a greater incumbency advantage for all statewide offices, just as reported by Ansolabehere and Snyder (2002).

2 The Model

At each of the three dates $t = 0, 1, 2$, a voter must delegate some responsibility to a representative. Each candidate for the job has ability $\theta \in \mathbb{R}$ which are independent draws from a

$\mathcal{N}(0, \sigma_\theta^2)$ distribution. The voter prefers candidates with high ability.

At the beginning of date 0, there is an incumbent in office. At the end of periods 0 and 1, the voter has the option of replacing the incumbent with a challenger. Each challenger is a fresh draw from the ability distribution. (We could allow for an election before date 0, but the voter would have no information about the candidates, so nothing would change.)

A complication enters the voter’s decision problem because the θ are unobserved. The only information the voter has is a signal, $s = \theta + \epsilon$, that is observed each period the candidate holds office. The shock ϵ is distributed $\mathcal{N}(0, \sigma_\epsilon^2)$, and is independent of θ .

In addition to the ability, each candidate has an ideological position. In particular, there are a left (L) and a right (R) party, with fixed policy platforms in the one-dimensional policy space. We denote these platforms by $\mu_p \in \mathbb{R}$, where $p \in \{L, R\}$. To keep the model as symmetric as possible, we assume that $\mu_L = -\mu_R$.² In each election, each party runs a candidate who is a perfect agent of the party. The incumbent is always that candidate for her party.

The representative voter has preferences represented by $-(x^* - x)^2 + \theta$, where x is a policy and x^* is the voter’s ideal point. The candidates do not know x^* ; their common belief is that $x^* \sim \mathcal{N}(\gamma, \sigma_{x^*}^2)$.

We assume that the voter acts myopically, always electing the candidate who offers the greatest next period utility. Notice that this implies the voter is risk neutral over ability. Unlike many papers in the literature, the incumbency advantage we study here has nothing to do with risk aversion—it is driven entirely by selection effects.

3 The Voter’s Decision Rule

Given the information structure outlined above, standard results (DeGroot, 1970) imply that the voter’s posterior belief about an incumbent’s ability, given a prior mean of m and a new signal s , are normal with mean

$$m' = \lambda s + (1 - \lambda)m, \tag{1}$$

²In our companion paper, candidates are not perfect agents of parties, so the voter believes that a party p ’s candidate’s policy preference is a random variable. To make the link between the papers, interpret μ_p as the voter’s certainty equivalent of this random variable.

where

$$\lambda = \frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_\epsilon^2}.$$

(We use m rather than 0 for the prior mean because the formula can be used to update after the second signal.) The posterior variance is equal to

$$\frac{\sigma_\theta^2 \sigma_\epsilon^2}{\sigma_\theta^2 + \sigma_\epsilon^2}.$$

The voter chooses which candidate to select by comparing the expected utility of each choice.

The voter votes for L if and only if

$$\mathbb{E}(-(\mu_L - x^*)^2 + \theta_L) \geq \mathbb{E}(-(\mu_R - x^*)^2 + \theta_R).$$

Taking the expectations and rearranging terms shows that this is true if

$$-(\mu_L - x^*)^2 + m_L \geq -(\mu_R - x^*)^2 + m_R,$$

which can be further reduced to

$$x^* \leq \frac{1}{2}(\mu_R + \mu_L) + \frac{m_L - m_R}{2(\mu_R - \mu_L)},$$

where m_c is the voter's expectation of candidate c 's ability. Since $\mu_L = -\mu_R$, this simplifies to

$$x^* \leq \frac{m_L - m_R}{2(\mu_R - \mu_L)}.$$

Without loss of generality, assume that the left-wing candidate is the incumbent. This incumbent is reelected if and only if:

$$x^* \leq \frac{m'_L(s) - m_R}{2(\mu_R - \mu_L)}$$

or

$$m'_L(s) \geq 2(\mu_R - \mu_L)x^* + m_R. \tag{2}$$

Define $\eta = 2(\mu_R - \mu_L)x^*$. Then η is normally distributed with mean 0 and variance $\sigma_\eta^2 = 4(\mu_R - \mu_L)^2 \sigma_{x^*}^2$. Since the voter's prior belief about the challenger's ability has mean $m_R = 0$ we can combine equations (1) and (2) to implicitly define the voter's re-election rule by:

$$\lambda s + (1 - \lambda)m > \eta.$$

With this reelection rule in hand, we are ready to turn to our main task, calculating the incumbency advantage and seeing what happens to it as the electoral environment changes.

4 Comparative Statics for a “Swing” District

We will start our comparative statics analysis by focusing on a “swing” district—one with $\gamma = 0$. Substantively, this means the district is equally likely to elect either party in an open seat election. Focusing on this case will allow us build intuition for how the model works in the simplest case, where we can get exact analytical results. The next section studies the general case, and uses a mix of analytical results and simulations to study the comparative statics with respect to competitiveness (γ), and to explore the robustness of the results in this section.

Before we derive the comparative statics, we need to be more precise about what the incumbency advantage is. Since all candidates are *ex-ante* identical and we are considering a swing district, the candidates in an open-seat election would each win with probability $1/2$. The incumbency advantage, then, is the extent to which an incumbent’s probability of defeating a challenger, who has never served in office, is greater than $1/2$.

The expected value of the posterior mean ability of the date 0 incumbent is 0—otherwise the law of iterated expectations would be violated. Combined with the normality of the prior distribution of the posterior mean, this implies that the date 0 incumbent wins reelection with probability $1/2$. Thus there is no incumbency advantage in the date 0 election. Things are different at date 1. If the date 1 incumbent was a challenger in the date 0 election, then the same argument as before shows that she wins at date 1 with conditional probability $1/2$. This event itself has probability $1/2$. If the date 1 incumbent won the date 0 election as an incumbent, however, things are different. Winning the first election must have been the result of a good signal, so this incumbent is likely better than average. (This is similar to the selection result in Gowrisankaran et al. (2003).)

Proposition 1 *In the date 1 election, the incumbent wins with probability strictly greater than $1/2$.*

The proof of this and all subsequent propositions are in the appendix.

This incumbency advantage arises because a two-term incumbent is more likely to win than to lose. This is intuitive—a candidate gets to be a two-term incumbent precisely because she impressed the voter in her first term. This above average performance means that the voter’s

best estimate of her ability is some $m > 0$. An outside observer will not be able to figure out exactly what the voter's estimate is, but we can infer that it is positive.

The key step in establishing the existence of an incumbency advantage is deriving the probability a two-term incumbent wins. This is equal to

$$\int f(m) \left(1 - \Phi \left(\frac{-m}{\sigma} \right) \right) dm,$$

where f is the density of expected abilities for candidates who have won reelection once, and $\sigma = \sqrt{\lambda\sigma_\theta^2 + \sigma_\eta^2}$. To see the intuition for this representation, notice that $1 - \Phi \left(\frac{-m}{\sigma} \right)$ is the probability a two-term incumbent with expected ability m wins reelection. The integral averages this probability over the distribution of expected abilities for incumbents. The existence of the incumbency advantage follows from the fact that f first-order stochastically dominates the distribution of abilities for untested incumbents.³

Remark The extra probability of winning associated with being an incumbent is not the only way to define the incumbency advantage. Indeed, empirical work often estimates the advantage as the difference in the incumbent's share of the two-party vote and the incumbent's party's share of the two-party vote in open seat elections. Our results also apply to this definition, at least in a simple one-dimensional model of voter heterogeneity. Assume that there are a continuum of voters, with ideologies x_i distributed $\mathcal{N}(0, \sigma_x^2)$. Each voter's payoff is $-(\mu_p - x_i - \xi)^2 + m_p$, where $\xi \sim \mathcal{N}(\gamma, \sigma_\xi^2)$. In this model, there will be a cutpoint \bar{x} such that all voters with ideal points

$$x_i + \xi \leq \bar{x} = \frac{m_L - m_R}{2(\mu_R - \mu_L)}.$$

This is the same inequality we derived in section 3, with x^* replaced by $x_i + \xi$. Thus, in the special case where $\text{var}(x_i + \xi) = \sigma_{x^*}^2$, the fraction of voters who vote for L is the same as the probability that L wins in our main model. Of course, this is a knife-edge condition, and the two measures will typically not be identical. Nonetheless, even when they are not identical (which they are not in the data) the isomorphism between the two models shows that the comparative statics will be the same.

³See Mas-Collel, Whiston, and Green (1995, ch. 2) for a review of stochastic dominance.

4.1 More Precise Information

Our first comparative static result concerns variation in the informativeness of the voter's signal about the incumbent's ability. The smaller is the variance of the error term, the more information the voter gets in any period. The incumbency advantage arises precisely because the voter is becoming rationally conservative about replacing representatives who have proved themselves in the past. Naturally, the more information is revealed in any given round, the more certain the voter is in his high estimate of his representative, and so the more rationally conservative the voter becomes. Thus the smaller the variance of the error term, the stronger is the incumbency advantage.

A more precise intuition is as follows: A less noisy signal of the incumbent's ability means the voter gets better information about the incumbent. This translates into a mean-preserving spread of the distribution of posterior means. The election then truncates this distribution, throwing away the left tail. Spreading out a distribution and then throwing out the bad parts leaves us with only the spread out good part, which is an improvement. Formally, the distribution of posteriors given reelection improves in the first-order stochastic sense when the variance of the noise term decreases.

There is another, reinforcing, effect. When the voter gets better information, he becomes relatively confident more quickly. This leads him to be less responsive to subsequent signals. This means that, for each value of the posterior mean m , his assessment of the ability is less likely to drop below his outside option of 0. These two effects combine to generate the improvement in the incumbency advantage. Putting these effects together, we have the following result.

Proposition 2 *Assume that $\sigma_{x^*}^2 = 0$. The greater is the variance of the voter's signal (σ_ϵ^2), the smaller is the incumbency advantage.*

Although our analytical result is limited to the case where there is no uncertainty about the voter's ideal point, our simulations in the next section show that the result does not actually rely on that restriction.

There are several applications of this result. First, it allows us to rank incumbency advantages based on the visibility of the office. Some offices, such as Governor or U.S. Senator, attract more media attention than others, like state auditor. The natural way to model this

extra information is by a reduced signal variance. Thus proposition 2 predicts that more visible offices will have greater incumbency advantages, a pattern observed by Ansolabehere and Snyder (2002). Second, the result lets us compare single-member and multi-member districts. This theme is developed next.

4.1.1 Application: Multi-Member Districts and Divided Delegations

A single-member district (SMD) corresponds exactly to the model outlined in the previous section. Thus the voter reelects the incumbent if and only if

$$\lambda^{SMD}s + (1 - \lambda^{SMD})m \geq 0,$$

where $\lambda^{SMD} = \sigma_\theta^2 / (\sigma_\theta^2 + \sigma_\epsilon^2)$.

In a multi-member district, presidential system (MMD), the setup is exactly as above, except there are two representatives, $i = 1, 2$. Each has ability θ_i , and these abilities are independent. The voter observes a signal $s = \theta_1 + \theta_2 + \epsilon$. In order to assess the ability of a legislator i , the voter must filter out the ability of legislator $-i$. As such, legislator $-i$'s contribution can be treated as an additional error term in the voter's signal extraction problem with respect to legislator i . That is, the signal is noisier because there are two legislators participating. Thus the voter learns less each round in an MMD system and so becomes less rationally conservative, which weakens the incumbency advantage.

Corollary 1 *The date 1 incumbency advantage is greater in SMD than MMD systems.*

Divided delegations, in which the voters select one representative from each political party, are a common political outcome in multi-member electoral systems. The election of a divided delegation has important informational consequences which can effect the strength of the incumbency advantage. Elsewhere, we discuss the strategic dynamics that might lead voters to elect a divided delegation (Ashworth and Bueno de Mesquita, 2003a). In particular, divided delegations may help voters determine which of their representatives to credit for local public goods, thereby forcing legislators to work harder for their constituents. Here it is sufficient to introduce a simple version of this model to demonstrate its implications for the incumbency advantage.

Assume that there are two different types of local public goods (LPGs): conservative and liberal. Legislators affiliated with party R produce conservative LPGs while legislators affiliated with party L produce liberal LPGs. These LPGs are the voter's only signal about the incumbent's ability. For instance, in the United States, Republicans may be better able to procure political pork reflecting industry interests, while Democrats may be more successful procuring political pork that benefits labor unions. Indeed, even the same LPG can be directed in conservative or liberal directions. For instance, all politicians may agree that they would like a piece of federally protected land in their district (such as a national park or national forest). However, some politicians would like the use of that piece of land to be directed towards logging while other politicians would prefer to restrict access to hikers or hunters. The key point is that local public goods have an ideological component.

Call the amount of conservative LPGs produced y^C and the amount of liberal LPGs produced y^L . If the voter chooses a divided delegation, she receives two signals:

$$\begin{aligned}y^C &= \theta^R + \epsilon \\y^L &= \theta^L + \epsilon.\end{aligned}$$

If the voter chooses a unified delegation from the left-wing party she receives only one signal (the right-wing party case is analogous):

$$y^L = \theta_i + \theta_j + \epsilon.$$

In a divided delegation the voter can properly assess the credit to give to each legislator (other than that due to the stochastic element of the production function) just as in an SMD because the two different types of LPGs are distinguishable. The amount of conservative LPGs constitute a signal of legislator R 's ability while the amount of liberal LPGs constitute a signal of legislator L 's ability. Choosing a divided delegation restores the informational environment within an MMD to something resembling an SMD. Proposition (2), thus, implies the following result.

Corollary 2 *The date 1 incumbency advantage is larger in divided than in unified delegations*

This result is consistent with the empirical finding by Alesina et al. (1991) that the incumbency advantage in divided delegations in the Senate is stronger than in unified delegations.

4.2 More Variable Voter Ideologies

Returning to our main model, our second comparative static result concerns changes in the uncertainty associated with the voter's ideal point. Recall that an incumbent from the L party is reelected if and only if $\lambda s + (1 - \lambda)m \geq \eta$, where η is a stochastic term that represents the voter's ideological leanings. As the voter puts more weight on ideology and less on ability, the variance of this stochastic threshold increases. (An example that shows how this works is in the application below.) When η is more variable, the voter is more willing to reelect a slightly sub-par incumbent to keep his favored party in power. Similarly, he is willing to vote out a slightly above average incumbent if he does not like that incumbent's party. This will reduce the rate at which the distribution of incumbent abilities improves. This leads to the following result.

Proposition 3 *The greater is the variance of the ideological term (σ_η^2), the smaller is the incumbency advantage.*

Now we use this result to show that the selection model can explain the comparative evidence on incumbency advantages.

4.2.1 Application: Presidential vs. Parliamentary Systems

The main result is

Corollary 3 *The date 1 incumbency advantage is greater in single member district presidential systems than in single member district parliamentary systems.*

To show this formally, we must be specific about the electoral environment and the voter's optimal voting rule. The rest of this subsection derives this rule, using a variant of the model we use to explain variations in constituency service and party discipline in Ashworth and Bueno de Mesquita (2003a) and Ashworth and Bueno de Mesquita (2003b). We show that the variance of the ideological term is greater in a single member district parliamentary system (from now on we refer to this simply as a parliamentary system) than in a single member district presidential (presidential) system, holding constant the variance of the voter's ideal point.

There are n legislative districts, each of which works as above. The ideal points, x^* , are independent across districts.

Rather than attempt to model of the complex politics of bargaining in a legislature, we follow Grossman and Helpman (1999) by adopting a reduced form representation. In particular, we assume that national policy is determined by the average of the individual legislator's ideal points. This is denoted x_L .

In the Westminster system national policy is fully determined by the decision of the legislature. In a presidential system, the legislature must bargain with the President to determine national policy. We will work with a reduced form description of this bargaining as well.⁴ Following Alesina and Rosenthal (1995), we assume that policy in a presidential system is a weighted average of the legislature's proposal and the President's ideal point, with weight β on the legislative proposal: $x = \beta x_L + (1 - \beta)x_P$.

The voter chooses which candidate to select in round 2 by comparing the expected utility of each choice. The analysis of section 3 shows that the voter reelects the incumbent if and only if

$$\lambda(s) + (1 - \lambda)m > \eta^{Parl},$$

where $\eta^{Parl} = 2(\mu_R - \mu_L)x^*$. The voter's re-election rule in a presidential system is similar. However, the diminished affect of the legislature on national policy due to the role of the president affects the voter's calculus. In particular, in an parliamentary system, the incumbent is re-elected given a level of LPG provision (s) if and only if:

$$x^* \leq \frac{m'_L(s) - m_R}{2\beta(\mu_R - \mu_L)}$$

or

$$m'_L(s) \geq 2\beta(\mu_R - \mu_L)x^* + m_R. \tag{3}$$

We define $\eta^{Pres} = 2\beta(\mu_R - \mu_L)x^*$. Then η^{Pres} is normally distributed with mean 0 and variance $\sigma_{\eta^{Pres}}^2 = 4\beta^2(\mu_R - \mu_L)^2\sigma_{x^*}^2$. Again recalling that the voter's prior belief about the challenger's ability has mean $m_R = 0$ we can combine equations (1) and (3) to implicitly define voter's re-election rule in a presidential system by:

$$\lambda(s) + (1 - \lambda)m > \eta^{Pres}.$$

⁴In Ashworth and Bueno de Mesquita (2003b), we briefly discuss more detailed bargaining models.

Now we can see how the intuition of the beginning of this subsection is captured in the formal model. The greater variance of η^{Parl} reflects the greater influence of ideological considerations relative to incumbent ability in parliamentary systems than in presidential systems. This means that the parliamentary voter is more likely both to return a sub-par incumbent from his favored party and to replace a good incumbent from the wrong party. This will reduce the rate at which the distribution of incumbent abilities improves in parliamentary systems.

5 Changes in Competitiveness

Now we return to the most general form of our model, allowing districts to lean toward one of the two parties on average. In particular, we assume that the voter's ideal point is distributed $\mathcal{N}(\gamma, \sigma_{x^*}^2)$. A negative value of γ indicates a district which leans left, while a positive value indicates a district that leans right. A one-term L incumbent wins with probability $1 - \Phi(\gamma/\sigma)$ (an R incumbent wins with complementary probability), where σ variance of the one-signal posterior mean. We will use the probability of a one-term R incumbent winning an election as our measure of the *normal vote*. The incumbency advantage will be the difference between the probability a two-term incumbent wins and this normal vote.

In general, the incumbency advantage within a district will differ by party. Writing $\mathcal{IA}_p(\gamma)$ for the incumbency advantage of a two-term incumbent from party p in a district with average ideal point γ , we can write the overall incumbency advantage as

$$\mathcal{IA}(\gamma) = (1 - \Phi(\gamma/\sigma))\mathcal{IA}_L(\gamma) + \Phi(\gamma/\sigma)\mathcal{IA}_R(\gamma).$$

Unfortunately, we are unable to give a complete analytic treatment of the comparative statics of this incumbency advantage. Instead, we will derive some analytical results in a limiting case, to build intuition, and then use simulations to show that the conclusions of these analytical exercises are robust.

To enable our analytical results, we will make two special assumptions. First, we assume that the voter's ideal point is deterministic: $\sigma_{x^*}^2 = 0$. This means that the left-wing candidate wins if and only if her ability advantage over the right-wing candidate is at least γ . Second, we assume that there is a random walk in the true ability: $\theta_{ct} = \theta_{c(t-1)} + v_t$, where $var(v) =$

$(\sigma_\theta^2)^2 / (\sigma_\theta^2 + \sigma_\epsilon^2)$. This implies that the variance of beliefs about ability does not decline with tenure.

Let's see what can be said about the incumbency advantage with these assumptions. The first assumption implies that the conditional density of posterior means for one-term incumbents who are reelected is

$$\frac{(1/\sigma)\phi(m/\sigma)}{1 - \Phi(\gamma/\sigma)}$$

for $m \geq \gamma$, and zero otherwise. Such an incumbent wins a third term if her two-signal posterior mean is also at least γ . This event has probability $1 - \Phi((\gamma - m)/\sigma')$, where σ' is the variance of the two-signal posterior mean. The second assumption implies that $\sigma' = \sigma$. Thus the probability a two-term L incumbent wins is

$$\int_\gamma^\infty \frac{(1/\sigma)\phi(m/\sigma)}{1 - \Phi(\gamma/\sigma)} \left(1 - \Phi\left(\frac{\gamma - m}{\sigma}\right)\right) dm,$$

and the L incumbency advantage is

$$\mathcal{IA}_L(\gamma) = \int_\gamma^\infty \frac{(1/\sigma)\phi(m/\sigma)}{1 - \Phi(\gamma/\sigma)} \left(1 - \Phi\left(\frac{\gamma - m}{\sigma}\right)\right) dm - 1 + \Phi(\gamma/\sigma).$$

Note that because the model is symmetric, the incumbency advantage is symmetric (that is, $\mathcal{IA}_L(\gamma) = \mathcal{IA}_R(-\gamma)$).

Consider two districts with biases $\gamma' < \gamma < 0$. We will show that the L incumbency advantage is greater in the more competitive district: $\mathcal{IA}_L(\gamma) > \mathcal{IA}_L(\gamma')$. This result is intuitive. If $\gamma \ll 0$, then the district leans heavily toward L candidates. This leads it to elect L candidates, even those with low valences. This means there is little censoring of the valence distribution for these candidates, and incumbents are very similar to open seat candidates. In a balanced district, on the other hand, incumbents are (on average) more different than open seat candidates.

More precisely, we can decompose the difference into three parts. First, there is the probability that a two term incumbent has posterior at least the more stringent cutoff, γ . This probability is greater for an incumbent who passed the more stringent hurdle. Second, there is the probability that the L candidate wins the last election, given that the first incumbent's first posterior was in (γ', γ) . In the more competitive district, this initial incumbent lost, so the current L candidate is a fresh draw from the prior distribution. Such a candidate is stronger

than the weak candidate who won the less competitive election. This effect also increases the incumbency advantage. Third, a two-term incumbent has an easier time (a lower hurdle) in the less competitive district. This works in the opposite direction of the two previous effects. The proof shows that the first effect outweighs the third.

Proposition 4 *The incumbency advantage for a left-wing candidate is increasing in γ on \mathbb{R}_- .*

What happens when $\gamma > 0$? The first effect we have identified, that an incumbent who has passed a more stringent hurdle must have higher ability, still works to make the advantage increasing in γ . However, the second effect, which is based on comparing the prior distribution to the prior conditioned on $[\gamma', \gamma]$, now works in the opposite direction. When $\gamma' > 0$, conditioning on the subinterval *increases* the expected ability of the incumbent. Since these incumbents win only in the *less* competitive districts, this makes the advantage tend to decrease with γ . In fact, it is easy to see that the second effect must dominate as $\gamma \rightarrow \infty$: When γ is very large, a L candidate will essentially never win, whether as a one-term or a two-term incumbent. Thus the L incumbency advantage will increase with γ until some value greater than zero, and then decrease.

The overall incumbency advantage in a district is a weighted average of the L and R incumbency advantages, with the weight on L given by the normal vote. This means that most of the weight is on L for $\gamma \ll 0$, and most of the weight is on R for $\gamma \gg 0$. Since the L advantage is increasing when $\gamma \ll 0$ and the R advantage is decreasing when $\gamma \gg 0$, this suggests that the incumbency advantage is smallest when γ is far from 0, and it is greatest near $\gamma = 0$.

Examination of the proof of proposition 4 shows that our proof strategy does rely on our two special assumptions. In particular, the absence of uncertainty in the voter's ideology allows us to derive the exact functional form of the distribution of posterior means, and we use this functional form to show that the first effect identified prior to the statement of the result outweighs the third effect. Also, without the fact that $\sigma' = \sigma$, the second effect would have an ambiguous sign, since two normal random variables with different variances cannot be ranked by stochastic dominance. In the next subsection, we use simulations to test the robustness of the results.

5.1 Simulation Results

Because we are unable to derive more general results analytically, we use simulations to test robustness.⁵ In these simulations we explore a model that relaxes the two special assumptions made above. First, we allow there to be uncertainty about the voter’s ideal point ($\sigma_\eta^2 \neq 0$). Second, we do not impose stationarity on the variance of beliefs about ability. We report the results in figures (1) and (2).

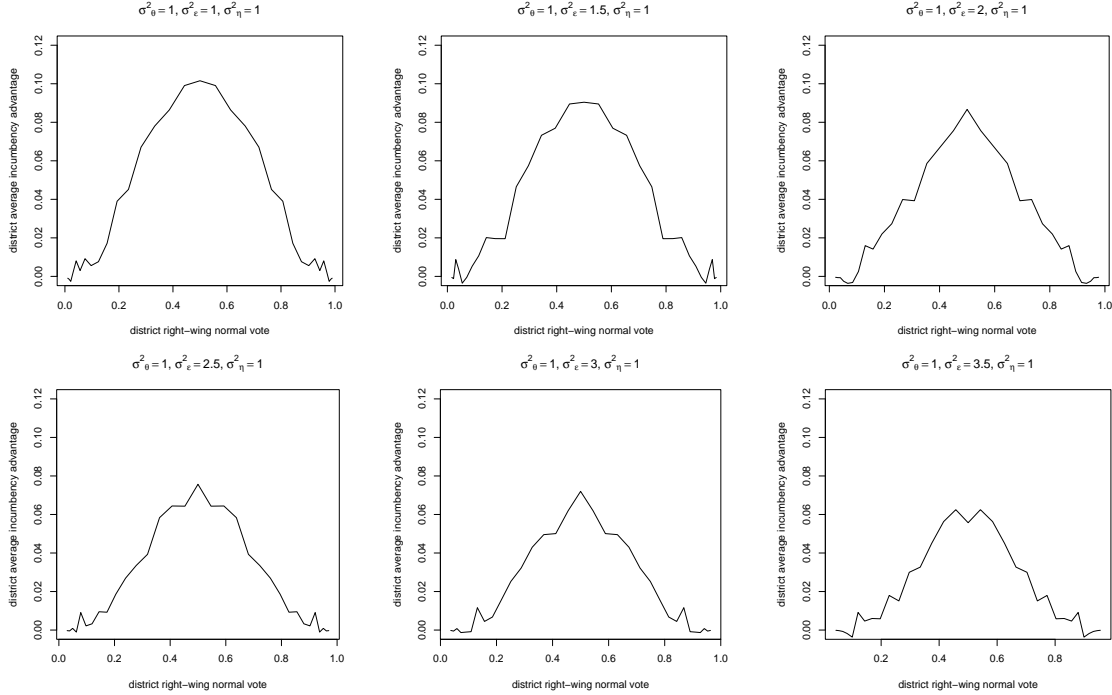


Figure 1: Comparative statics on σ_ϵ^2 and γ

Within this more general model, we investigate three comparative statics. First, we show that for a variety of parameter values, the comparative statics regarding changes in competitiveness (γ) are almost always true. In particular, the incumbency advantage is increasing as a district’s normal vote moves closer to 1/2 for all but the most extreme districts. Recall that the cause of potential non-monotonicity with respect to competitiveness had to do with a left-wing incumbent who holds office in a very right-wing district (or vice-versa). Thus, the only place where the non-monotonicity can be seen in the simulations is at districts so far to the extreme of the ideological (right-wing normal votes under 10 percent or over 90 percent) that they are

⁵The *R* code used in these simulations is available at www.artsci.wustl.edu/~ebuenode.

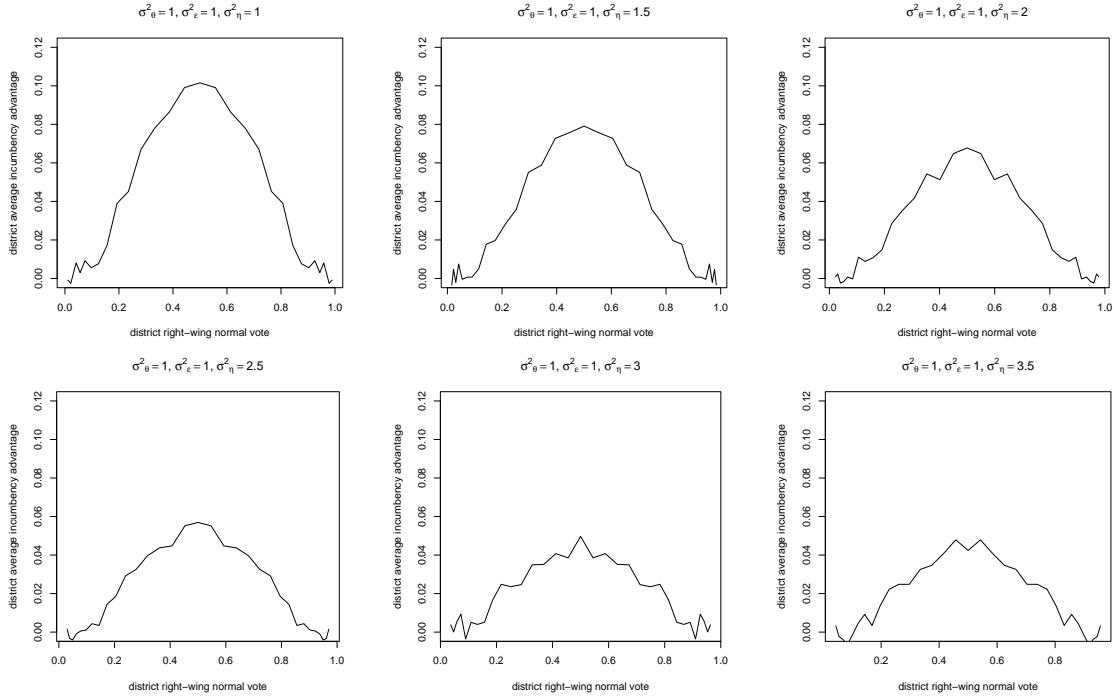


Figure 2: Comparative statics on σ_η^2 and γ

of no empirical interest. These results can be seen in each individual frame within both figures (1) and (2). Each frame represents a vector of parameter values $(\sigma_\theta^2, \sigma_\epsilon^2, \sigma_\eta^2)$. The x -axis is the normal vote for a right-wing candidate in a district. Thus, moving along the x -axis is equivalent to changing the level of district competitiveness (γ), where competitiveness is maximized at the midpoint. As expected, except at the very tails, for all vectors of parameter values that we have investigated, the incumbency advantage (the y -axis) is increasing as competitiveness increases.

We also check the robustness of our comparative statics on the variance of the shock to the signal (σ_ϵ^2) and the variance of the ideological component of the voter's decision rule (σ_η^2). In both cases the comparative statics in the simulated, more general model are the same as the analytically derived results from the simpler model. When the noise associated with the signal (σ_ϵ^2) increases, the incumbency advantage decreases. This is because more noisy signals lead to less censoring with each election. This can be seen by moving across the panels in figure (1). Each subsequent panel has a higher σ_ϵ^2 and, consequently, a lower incumbency advantage.

Similarly, greater variance in the ideological component of the voter's decision rule (σ_η^2) means that ideological considerations loom large. This leads voters to return somewhat sub-

par incumbents if they are of the preferred party and throw out somewhat above-par incumbents if they are not of the preferred party. This focus on ideology means that voters are not paying as much attention to the ability dimension (which is the source of the incumbency advantage) when they select candidates. Thus, as σ_{η}^2 increases, the incumbency advantage decreases. This can be seen by following the panels of figure (2).

5.2 Application: Change in the U.S. Incumbency Advantage Over Time

Now we argue that our results on the comparative statics of competitiveness can help explain the changes in the incumbency advantage over U.S. history. The main fact to be explained is that for all offices, the average incumbency advantage was small to nonexistent at mid-century, grew dramatically in the 1960s and 1970s, and has stabilized at a high level since then.

The selection model suggests the following explanation. At mid-century, most states were relatively uncompetitive. The south was solidly Democratic, while the non-urban parts of the north and west were solidly Republican. This trend reversed in the mid-1960s, due in part to the debate over civil rights. Today the parties are more competitive in most states. Erikson, Wright, and McIver (1993) find that public opinion is roughly balanced between the two parties in most states, and Ansolabehere and Snyder (2002) find that the portion of vote shares explained by the state's partisan leanings ("the normal vote") has declined dramatically since mid-century. They write:

The normal vote accounts for 53 percent of the variation in the vote in the 1940s. It's importance drops substantially in the 1950s, to 40 percent of total variance in the vote. And it collapses in the 1960s, explaining only 20 percent of the variance in the vote in the 1960s and 1970s. The decline of the normal vote as an explanatory factor continues in the 1980s, falling to 10 percent in the 1980s and 1990s.

In terms of the model, the "solid south" is a group of districts with $\gamma \ll 0$, while most of the north and west had $\gamma \gg 0$. Both of these lead to small incumbency advantages. Today most states have $\gamma \approx 0$, which leads to a greater incumbency advantage.

6 Conclusion

We have shown that the electoral selection view can explain the main comparative facts about the incumbency advantage. Since the selection theory is based on the flow of information to the voter, variations in the incumbency advantage are driven by variation in either the quality of the information or the willingness of the voter to make use of the information.

Although this paper shows that the electoral selection view can account for the comparative and historical facts about the incumbency advantage, it does not address the empirically established links between constituency service and incumbent success. We address this in other work. Ashworth (2001) studies a model with both symmetric learning and moral hazard. Learning leads to an incumbency advantage, just as in this paper. In addition, incumbents use constituency service activity to try and convince the voter that they have high ability. Variation in constituency service leads to variation in incumbent success, but it is a fallacy of composition to argue that constituency service causes the average level of incumbency advantage. In fact, given the specific assumptions of that paper, the incumbency advantage would be the same if even if constituency service were stopped entirely. Finally, Ashworth and Bueno de Mesquita (2003a) show that the same factors that lead to variation in the incumbency advantage in this paper lead to similar variations in constituency service. Thus the selection model, augmented with a moral hazard dimension, can explain the joint comparative statics of incumbency advantage and constituency service, and can do so while treating both as endogenous.

A Proofs

At each election, the voter's beliefs about the incumbent's ability will be normally distributed with mean m and variance σ_n^2 , where n is the number of signals the voter has observed about that incumbent's ability. By assumption, $\sigma_0^2 = \sigma_\theta^2$, and a standard result tells us that

$$\sigma_n^2 = \frac{\sigma_{n-1}^2 \sigma_\epsilon^2}{\sigma_{n-1}^2 + \sigma_\epsilon^2}.$$

We can often simplify notation by defining

$$\lambda_n = \frac{\sigma_n^2}{\sigma_n^2 + \sigma_\epsilon^2}.$$

We will drop the subscript on λ when no confusion will result.

We let f denote the density of posterior means for candidates who have won reelection once. After this candidate's performance in her second term, the voter's belief can be written as

$$m' = m + \lambda_1(\delta + \epsilon),$$

where $\delta = \theta - m$ is the error in the date 0 beliefs. A standard result tells us that δ and m are independent random variables. This incumbent wins another reelection if and only if $m' - \eta \geq \gamma$.

A.1 Preliminary Results

We begin with some preliminary results that will be useful in all of the proofs. Denote the standard normal cdf by Φ and its density by ϕ , and let $\sigma = \sqrt{\lambda_1 \sigma_1^2 + \sigma_\eta^2}$.

Lemma 1 *A two-term incumbent wins reelection with probability*

$$\int f(m) \left(1 - \Phi \left(\frac{\gamma - m}{\sigma} \right) \right) dm.$$

Proof A standard result is that, if x and y are independent random variables with densities f and g , respectively, then $x + y$ has density $h(x) = \int f(y)g(x - y) dy$. We use this result to find the density of $m' - \eta = m + \lambda(\delta + \epsilon) - \eta$. Here, m is a constant, $\lambda(\delta + \epsilon)$ is distributed $\mathcal{N}(0, \lambda\sigma_\theta^2)$, and η is distributed $\mathcal{N}(\gamma, \sigma_\eta^2)$. Thus the difference $\lambda(\delta + \epsilon) - \eta$ is also normally distributed, with mean $-\gamma$ and variance $\lambda\sigma_1^2 + \sigma_\eta^2$. Thus the density of $m' - \eta$ is

$$h(y) = \int f(m) \frac{1}{\sigma} \phi \left(\frac{y - m + \gamma}{\sigma} \right) dm.$$

The probability of reelection is the probability that $m' - \eta > 0$. To find this we need the decumulated distribution function, which we find by integrating:

$$1 - H(y) = \int f(m) \left(1 - \Phi \left(\frac{y + \gamma - m}{\sigma} \right) \right) dm.$$

Evaluating this at $y = 0$ gives the result. \square

Now we prove a couple of comparative static results about this probability. Let r denote the electoral regime under consideration, and let f_r be the density of date 1 posteriors given the equilibrium voting rule at the first election in regime r .

Lemma 2 *If $f_1 \succ_{FOSD} f_2$, then the reelection probability is greater in regime 1 than in regime 2.*

Proof This follows from a standard result on stochastic dominance, since $1 - \Phi((\gamma - m)/\sigma)$ is increasing in m . \square

Lemma 3 *Assume that $\gamma = 0$ and that $f(x) > f(-x)$ for all $x > 0$. Then the reelection probability is decreasing in σ .*

Proof Differentiate the reelection probability to get

$$\begin{aligned} \frac{d}{d\hat{\sigma}} \text{Pr} &= \int -\phi \left(\frac{-m}{\hat{\sigma}} \right) \cdot \frac{m}{\hat{\sigma}^2} f(m) dm \\ &= \frac{-1}{\hat{\sigma}^2} \left(\int_{-\infty}^0 m\phi(-m/\hat{\sigma})f(m) dm + \int_0^{\infty} m\phi(-m/\hat{\sigma})f(m) dm \right). \end{aligned}$$

In the last line, the first integrand is negative, while the second is positive. Thus we will be done as soon as we show that the absolute value of the second integral is greater than the absolute value of the first. We have

$$\begin{aligned} \left| \int_{-\infty}^0 m\phi(-m/\hat{\sigma})f(m) dm \right| &< \int_{-\infty}^0 |m\phi(-m/\hat{\sigma})f(m)| dm \\ &= \int_{-\infty}^0 -m\phi(-m/\hat{\sigma})f(m) dm \\ &= \int_0^{\infty} m\phi(m/\hat{\sigma})f(-m) dm \\ &= \int_0^{\infty} m\phi(-m/\hat{\sigma})f(-m) dm \\ &< \int_0^{\infty} m\phi(-m/\hat{\sigma})f(m) dm, \end{aligned}$$

where the first equality is the definition of the absolute value, the third equality is from the symmetry of the normal density, and the last inequality is from the hypothesis on f . \square

A.2 Proof of Proposition 1

First, consider the reelection probability at the second election under the (counterfactual) assumption that all incumbents are reelected at the first election. In this case, the distribution of posteriors is $\Phi(m/\sigma_m)$, and the reelection probability at the second election is

$$\begin{aligned}
Prob &= \int \phi\left(\frac{m}{\sigma_m}\right) \left(1 - \Phi\left(\frac{-m}{\sigma_m}\right)\right) dm \\
&= \int_0^\infty \phi\left(\frac{-m}{\sigma_m}\right) \left(1 - \Phi\left(\frac{m}{\sigma_m}\right)\right) dm + \int_0^\infty \phi\left(\frac{m}{\sigma_m}\right) \left(1 - \Phi\left(\frac{-m}{\sigma_m}\right)\right) dm \\
&= \int_0^\infty \phi\left(\frac{m}{\sigma_m}\right) \Phi\left(\frac{-m}{\sigma_m}\right) dm + \int_0^\infty \phi\left(\frac{m}{\sigma_m}\right) \left(1 - \Phi\left(\frac{-m}{\sigma_m}\right)\right) dm \\
&= \int_0^\infty \phi\left(\frac{m}{\sigma_m}\right) dm \\
&= \frac{1}{2},
\end{aligned}$$

where the second equality is a change of variables and the third equality follows from the symmetry of the normal distribution.

The actual density of posteriors is $f(m) \propto \phi(m/\sigma_m) (1 - \Phi(-m/\sigma_\eta))$. Thus if we fix $m > m'$, the likelihood ratio is

$$\frac{\phi(m/\sigma_m) (1 - \Phi(-m/\sigma_\eta))}{\phi(m'/\sigma_m) (1 - \Phi(-m'/\sigma_\eta))} > \frac{\phi(m/\sigma_m)}{\phi(m'/\sigma_m)},$$

where the inequality follows from monotonicity of the normal cdf. Thus the actual distribution of posteriors MLR-dominates the no selection distribution we considered above, and, since the MLR order is a subset of the FOSD order, lemma 2 implies the proposition.

A.3 Proof of Proposition 2

To prove the proposition, we appeal to the previous results in the special case where both γ and σ_η^2 are 0. In this case, the incumbent is reelected in the first round if and only if the posterior mean was at least 0. Thus m has a truncated normal distribution, with density $(2/\sigma_m)\phi(m/\sigma_m)$.

Let $\sigma_m^2 = \lambda_0 \sigma_0^2$ be the variance of the date 0 posterior mean, and let $\sigma_1^2 = \lambda_0 \sigma_\epsilon^2$ be the date 0 posterior variance. Then the reelection probability can be written as

$$2 \int_0^\infty \frac{1}{\sigma_m} \phi\left(\frac{m}{\sigma_m}\right) \left(1 - \Phi\left(\frac{-m}{\sqrt{\lambda} \sigma_1}\right)\right) dm.$$

We claim that $(2/\sigma_m)\phi(m/\sigma_m)$ is increasing (in the FOSD sense) in σ_m and that $\lambda\sigma_1^2$ is decreasing in σ_ϵ^2 . To see that these facts prove the proposition, let regime S have a smaller value of σ_ϵ^2 than regime M , and notice that

$$\begin{aligned} \text{Inc adv M} &= 2 \int_0^\infty \frac{1}{\sigma_m^M} \phi\left(\frac{m}{\sigma_m^M}\right) \left(1 - \Phi\left(\frac{0-m}{\sqrt{\lambda} \sigma_1^M}\right)\right) dm \\ &< 2 \int_0^\infty \frac{1}{\sigma_m^S} \phi\left(\frac{m}{\sigma_m^S}\right) \left(1 - \Phi\left(\frac{0-m}{\sqrt{\lambda} \sigma_1^M}\right)\right) dm \\ &< 2 \int_0^\infty \frac{1}{\sigma_m^S} \phi\left(\frac{m}{\sigma_m^S}\right) \left(1 - \Phi\left(\frac{0-m}{\sqrt{\lambda} \sigma_1^S}\right)\right) dm \\ &= \text{Inc Adv S}, \end{aligned}$$

where the first inequality follows from $f_S \succ_{FOSD} f_M$ and lemma 2, and the second inequality follows from $\lambda\sigma_1^2$ is decreasing in σ_ϵ^2 and lemma 3.

Now we prove the claims. First consider the ranking of the density of posteriors.

Lemma 4 $\Phi\left(\frac{m}{\sigma}\right)$ is decreasing in σ if $m > 0$ and is increasing in σ if $m < 0$.

Proof Differentiate to get

$$\frac{\partial}{\partial \sigma} \Phi\left(\frac{m}{\sigma}\right) = \phi\left(\frac{m}{\sigma}\right) \left(-\frac{m}{\sigma^2}\right),$$

which has the same sign as $-m$. □

The cdf of the posteriors is 0 on $(-\infty, 0)$ and is $2(\Phi(m/\sigma_m) - 1/2)$ on $[0, \infty)$. Thus the claim follows from the fact that

$$\sigma_m^2 = \frac{\sigma_0^4}{\sigma_0^2 + \sigma_\epsilon^2}$$

is decreasing in σ_ϵ^2 .

Finally, we prove the second claim. Calculate

$$\begin{aligned} \frac{d}{d\sigma_\epsilon^2} \frac{\sigma_0^2 \sigma_\epsilon^2}{\sigma_0^2 + \sigma_\epsilon^2} &= \frac{\sigma_0^2(\sigma_0^2 + \sigma_\epsilon^2) - \sigma_0^2 \sigma_\epsilon^2}{(\sigma_0^2 + \sigma_\epsilon^2)^2} \\ &= \frac{\sigma_0^4}{(\sigma_0^2 + \sigma_\epsilon^2)^2} \\ &> 0, \end{aligned}$$

to see that σ_1^2 is increasing in σ_ϵ^2 . Then, using the definition of λ_1 , we have

$$\begin{aligned}\frac{d}{d\sigma_1^2}\lambda_1\sigma_1^2 &= \frac{2\sigma_1^2(\sigma_1^2 + \sigma_\epsilon^2) - \sigma_1^4}{(\sigma_1^2 + \sigma_\epsilon^2)^2} \\ &> 0.\end{aligned}$$

Together, these two results complete the proof.

A.4 Proof of Proposition ??

Now we turn to the comparison of the US and the UK. Let f_r be the density of the date 1 posterior means (given reelection at date 1) in system $r \in \{S, K\}$, where S has a smaller value of σ_η^2 than does K. Notice that, conditional on a date 0 posterior m , these densities are the same—they differ only in the likelihood of any particular date 0 posterior.

First, observe that the (preelection) density of date 0 posteriors is $(1/\sigma_m)\phi(\cdot/\sigma_m)$. The postelection density at x is thus proportional to $(1/\sigma_m)\phi(x/\sigma_m)(1 - \Phi(-x/\sigma_\eta))$. Since the normal density is symmetric about 0, we can rewrite this as $(1/\sigma_m)\phi(x/\sigma_m)\Phi(x/\sigma_\eta)$.

Lemma 5 *The density f is indexed by $-\sigma_\eta$ in the MLRP sense.*

Proof Fix $x > x'$. The likelihood ratio is

$$\mathcal{LR} = \frac{\phi(x/\sigma_m)\Phi(x/\sigma_\eta)}{\phi(x'/\sigma_m)\Phi(x'/\sigma_\eta)}.$$

We need to show that this is decreasing in σ_η . Differentiate to get

$$\frac{\partial \mathcal{LR}}{\partial \sigma_\eta} = \frac{\phi(x/\sigma_m)}{\phi(x'/\sigma_m)} \frac{(-x/\sigma_\eta^2)\phi(x/\sigma_\eta)\Phi(x'/\sigma_\eta) - (-x'/\sigma_\eta^2)\phi(x'/\sigma_\eta)\Phi(x/\sigma_\eta)}{\Phi(x'/\sigma_\eta)^2}.$$

This has the same sign as

$$(-x/\sigma_\eta^2)\phi(x/\sigma_\eta)\Phi(x'/\sigma_\eta) - (-x'/\sigma_\eta^2)\phi(x'/\sigma_\eta)\Phi(x/\sigma_\eta),$$

so $\partial \mathcal{LR}/\partial \sigma_\eta < 0$ if and only if

$$(-x/\sigma_\eta^2)\phi(x/\sigma_\eta)\Phi(x'/\sigma_\eta) < (-x'/\sigma_\eta^2)\phi(x'/\sigma_\eta)\Phi(x/\sigma_\eta)$$

if and only if

$$-x \frac{\phi(x/\sigma_\eta)}{\Phi(x/\sigma_\eta)} < -x' \frac{\phi(x'/\sigma_\eta)}{\Phi(x'/\sigma_\eta)}.$$

This inequality follows from $x > x'$ and the logconcavity of ϕ . \square

We will use the following result from Athey (2002): If $f(x, y)$ is logsupermodular in x and y and $g(y, z)$ is logsupermodular in y and z , then $h(x, z) = \int f(x, y)g(y, z) dy$ is logsupermodular in x and z .

Lemma 6 *For either r , $f_r(x) > f_r(-x)$.*

Proof If $x > x'$, we can use the monotonicity of the normal cdf to get

$$\begin{aligned} & \frac{\phi(x/\sigma_m)(1 - \Phi(-x/\sigma_\eta))}{\phi(x'/\sigma_m)(1 - \Phi(-x'/\sigma_\eta))} \\ & > \frac{\phi(x/\sigma_m)}{\phi(x'/\sigma_m)}, \end{aligned}$$

so the postelection date 0 posterior MLR dominates the preelection posterior distribution. Thus the lemma from Athey implies that the convolution of the postelection density with the posterior innovation density MLR dominates the convolution of the preelection density with the innovation density. But this second convolution yields a mean-zero normal density. Since this density is symmetric, the result follows. \square

Since the MLR order is a subset of the FOSD order and σ_η^2 is greater in K by hypothesis, lemmas 2 and 3 imply that the reelection probability is greater in S, proving the proposition.

A.5 Proof of Proposition 4

We have

$$\begin{aligned} \sigma (\mathcal{I}\mathcal{A}_L(\gamma) - \mathcal{I}\mathcal{A}_L(\gamma')) &= \int_\gamma^\infty \frac{\phi(m/\sigma)}{1 - \Phi(\gamma/\sigma)} (1 - \Phi((\gamma - m)/\sigma)) dm \\ &\quad - \int_\gamma^\infty \frac{\phi(m/\sigma)}{1 - \Phi(\gamma'/\sigma)} (1 - \Phi((\gamma' - m)/\sigma)) dm \\ &\quad - \int_{\gamma'}^\gamma \frac{\phi(m/\sigma)}{1 - \Phi(\gamma'/\sigma)} (1 - \Phi((\gamma' - m)/\sigma)) dm \\ &\quad + \int_{\gamma'}^\gamma \phi(m/\sigma) dm. \end{aligned}$$

Consider first the last two terms. Since $m \in (\gamma', \gamma)$ implies $\gamma' - m > \gamma'$, we have

$$\frac{1 - \Phi((\gamma' - m)/\sigma)}{1 - \Phi(\gamma'/\sigma)} < 1$$

for all $m \in (\gamma', \gamma)$. Thus we have

$$\int_{\gamma'}^{\gamma} \phi\left(\frac{m}{\sigma}\right) \left(1 - \frac{1 - \Phi((\gamma' - m)/\sigma)}{1 - \Phi(\gamma'/\sigma)}\right) dm > 0,$$

since the integrand is positive (almost surely).

Now we turn to the first two terms. The first term is

$$\frac{1}{1 - \Phi(\gamma/\sigma)} \int_{\gamma}^{\infty} \phi(m/\sigma)(1 - \Phi((\gamma - m)/\sigma)) dm.$$

Integrate this by parts to get

$$\begin{aligned} & \frac{1}{1 - \Phi(\gamma/\sigma)} \int_{\gamma}^{\infty} \phi(m/\sigma)(1 - \Phi((\gamma - m)/\sigma)) dm \\ &= \left(1 - \Phi\left(\frac{\gamma - m}{\sigma}\right)\right) \Phi\left(\frac{m}{\sigma}\right) \Big|_{\gamma}^{\infty} + \int_{\gamma}^{\infty} \Phi\left(\frac{m}{\sigma}\right) \phi\left(\frac{\gamma - m}{\sigma}\right) dm \\ &= 1 - \frac{1}{2} \Phi\left(\frac{\gamma}{\sigma}\right) + \int_{\gamma}^{\infty} \Phi\left(\frac{m}{\sigma}\right) \phi\left(\frac{\gamma - m}{\sigma}\right) dm. \end{aligned}$$

Similarly, integrate the second term by parts to get

$$1 - \left(1 - \Phi\left(\frac{\gamma' - \gamma}{\sigma}\right)\right) \Phi\left(\frac{\gamma}{\sigma}\right) + \int_{\gamma}^{\infty} \Phi\left(\frac{m}{\sigma}\right) \phi\left(\frac{\gamma' - m}{\sigma}\right) dm.$$

Since $1 - \Phi((\gamma' - \gamma)/\sigma) > 1/2$, the first term in the sum is greater in the first integral. Also, the density in the first integral dominates that in the second in the sense of FOSD and Φ is increasing, so the second term of the first integral is greater than the second term of the second integral. Finally, the prior probability that an L candidate wins is greater in the less competitive district, so the first integral is divided by a smaller number than is the second integral. Together, these imply that the first term in the decomposition dominates the second.

References

- Alesina, A., M. Fiorina, and H. Rosenthal (1991). Why are there so many divided Senate delegations? NBER Working Paper No. 3663.
- Alesina, A. and H. Rosenthal (1995). *Partisan Politics, Divided Government, and the Economy*. New York: Cambridge University Press.
- Ansola-behere, S. and J. M. Snyder, Jr. (2002). The incumbency advantage in U.S. elections: An analysis of state and federal offices, 1942–2000. *Election Law Journal* 1, 315–338.
- Ashworth, S. (2001). Reputational dynamics and political careers. Unpublished Paper, Harvard.
- Ashworth, S. and E. Bueno de Mesquita (2003a). Constituency service with electoral and institutional variation. Unpublished Paper.
- Ashworth, S. and E. Bueno de Mesquita (2003b). Party discipline with electoral institutional variation. Unpublished Paper.
- Athey, S. (2002). Monotone comparative statics under uncertainty. *Quarterly Journal of Economics* CXVII, 187–223.
- Cain, B., J. Ferejohn, and M. Fiorina (1987). *The Personal Vote: Constituency Service and Electoral Independence*. Harvard University Press.
- Cox, G. W. and J. N. Katz (2002). *Elbridge Gerry's Salamander: The Electoral Consequences of the Reapportionment Revolution*. New York: Cambridge University Press.
- Cox, G. W. and S. Morgenstern (1995). The incumbency advantage in multimember districts: Evidence from the U.S. states. *Legislative Studies Quarterly* 20, 329–349.
- DeGroot, M. H. (1970). *Optimal Statistical Decisions*. McGraw-Hill.
- Erikson, R. S., G. C. Wright, and J. P. McIver (1993). *Statehouse Democracy: Public Opinion and Policy in the American States*. Cambridge: Cambridge University Press.
- Fiorina, M. P. (1977). *Congress: Keystone of the Washington Establishment*. Yale University Press.

- Fiorina, M. P. and D. Rivers (1989). Constituency service, reputation, and the incumbency advantage. In M. P. Fiorina and D. Rohde (Eds.), *Home Style and Washington Work*. University of Michigan Press.
- Gelman, A. and G. King (1990). Estimating incumbency advantage without bias. *American Journal of Political Science* 34, 1142–1164.
- Gowrisankaran, G., M. F. Mitchell, and A. Moro (2003). Why do incumbent senators win? evidence from a dynamic selection model. Unpublished paper.
- Grossman, G. M. and E. Helpman (1999). Competing for endorsements. *American Economic Review* 89, 501–524.
- Katz, J. and G. King (1999). A statistical model for multiparty electoral data. *American Political Science Review* 93, 15–32.
- Katz, R. (1986). Interparty preference voting. In B. Grofman and A. Lijphart (Eds.), *Electoral Laws and their Political Consequences*. New York: Agathon Press.
- Levitt, S. D. and J. M. Snyder, Jr. (1997). The impact of federal spending on house election outcomes. *Journal of Political Economy* 105, 30–53.
- Levitt, S. D. and C. D. Wolfram (1997). Decomposing the sources of incumbency advantage in the U.S. house. *Legislative Studies Quarterly* 22, 45–60.
- Mas-Collel, A., M. D. Whiston, and J. R. Green (1995). *Microeconomic Theory*. Oxford: Oxford University Press.
- Zaller, J. (1998). Politicians as prize fighters: Electoral selection and the incumbency advantage. Unpublished Manuscript.