Francesco Caselli, Tom Cunningham, Massimo Morelli and Inés Moreno de Barreda

The incumbency effects of signalling

Article (Accepted version) (Refereed)

Original citation:
Caselli, Francesco, Cunningham, Tom, Morelli, Massimo and de Barreda, Inés Moreno (2014) *The incumbency effects of signalling*. Economica, 81 (323). pp. 397-418. ISSN 1468-0335
DOI: 10.1111/ecca.12060

The Incumbency Effects of Signalling[1]

Francesco Caselli, Tom Cunningham, Massimo Morelli and Inés Moreno de Barreda

January 2013

[1]Francesco Caselli: London School of Economics. E-mail: f.caselli@lse.ac.uk. Tom Cunningham: Harvard University. E-mail: tcunning@fas.harvard.edu. Massimo Morelli: Columbia University. E-mail: mm3331@columbia.edu. Inés Moreno de Barreda: University of Oxford. E-mail: ines.morenodebarreda@economics.ox.ac.uk. We thank, for helpful comments, Navin Kartik, Meg Meyer, Ronny Razin, Francesco Squintani, and seminar participants at LSE, Zurich, Oxford, Edinburgh, and at the 2007 Summer School on Political Economy and Social Choice. Previous versions of this paper circulated with the titles ≪Incumbency Advantage and Supermajority Reelection Rule≫, ≪Handicapping Politicians: the Optimal Majority Rule in Incumbency Elections≫ and ≪Signalling, Incumbency Advantage, and Optimal Reelection Rules≫.

Abstract

Much literature on political behavior treats politicians as motivated by reelection, choosing actions to signal their types to voters. We identify a novel implication of incumbent signalling. Because incumbents only care about clearing a reelection hurdle, signals will tend to cluster just above the threshold needed for reelection. This generates a skew distribution of signals leading to an incumbency advantage in the probability of election. We also solve for the optimal threshold when voters have the ability to commit.

JEL Classifications: D72, D78, D82.

Keywords: Signalling, Incumbency advantage, Supermajority.

1 Introduction

It has long been recognized that incumbent politicians can take actions in order to affect voters' perception of their types, and that such signalling has the potential to explain important empirical phenomena. For example, it has been argued that high-ability incumbents may have incentives to engineer pre-election booms in order to distinguish themselves from low-ability ones [Rogoff (1990), Rogoff and Sibert (1988)]. Signalling has also been used to explain "pandering", where the incumbent may have incentives to ignore private information and pander to the prior opinion of the median voter [e.g. Canes-Wrone et al. (2001), Maskin and Tirole (2004), Morelli and Van Weelden (2011a)], and "posturing", where the incumbent puts effort on a divisive issue that helps signal the congruence of his or her preferences with the majority [e.g. Morelli and Van Weelden (2011b), Acemoglu et al. (2011)].

This paper identifies a new important implication of signalling opportunities that are available only, or mainly, to the incumbent. As long as voters are rational, receiving a signal will not bias their evaluation of the incumbent on average. However because incumbents only care about clearing a reelection hurdle, signals will tend to cluster just above the threshold needed for reelection. This generates a skew distribution of signals leading to an incumbency advantage in the probability of election.

It is well known that incumbents enjoy disproportionate reelection rates, even in countries where the electoral process is generally deemed to be free and fair [Gelman and King (1990)]. In part, this phenomenon can be explained by selection: the pool of incumbents who run for reelection may contain a disproportionately large fraction of high-quality politicians. However, there is a broad consensus, recently buttressed by new empirical findings aiming at controlling for selection [Ansolabehere et al. (2000), Lee (2008), Levitt and Wolfram (1997)], that incumbency *per se* has a causal effect on election outcomes.[1] In other words incumbency confers direct electoral benefits even when candidates are of similar quality. This paper shows that such a causal incumbency advantage is a natural side-effect of the advantage that sitting politicians have in being able to signal their type.[2]

We present a model in which voters choose between an incumbent and a challenger, with both politicians drawn from the same symmetric distribution of three types: low,

[1] Levitt and Wolfram (1997) compare repeated pairings of candidates for election to the US Congress, in an attempt to control for the quality of incumbent and challenger. They find that the winner of the previous race has on average a 4% higher vote share in the second pairing. Ansolabehere et al. (2000) compare county-level vote shares after redistricting in US Congressional elections. They find that incumbents receive 4% fewer votes in counties which have been redistricted into their constituencies, than in counties which remained in their constituency for both elections. Lee (2008) compares bare winners and bare losers of elections. He finds that a party which barely wins a Congressional election has on average an 8% higher vote share and a 35% higher probability of winning the next election.

[2] An alternative explanation for a causal incumbency advantage is that incumbents "improve" with tenure, either thanks to the accumulated experience, or because seniority makes them more influential [Dick and Lott (1993)]. This explanation and our signalling explanation are not mutually exclusive.

middle, and high quality.[3] This signal has three components: the true quality of the incumbent, the signalling effort exerted by the incumbent, and noise.[45] Under fairly mild assumptions middling-quality incumbents exert the greatest signalling efforts, as high-quality ones can rely on their true quality to convince voters, while low-quality ones are discouraged by the excessive effort costs that would be required to fool voters. Therefore, the distribution of signals observed by voters is skewed: more than 50 percent of signals lead to posterior expected quality that is greater than the average quality (i.e., greater than the expected quality of the challenger). As a result, more than 50 percent of incumbents will be re-elected. This implies an incumbency advantage. In particular, middle-quality incumbents win elections against challengers (whose expected quality is middle) more than 50% of the time.

As already stated our paper contributes to the literature on signalling in games between politicians and voters. Most of this literature is concerned with how signalling incentives can explain some seemingly perverse policy outcome (e.g. an incentive to engineer a pre-election boom). Hence, it is of central importance for these studies' purposes that the signalling action itself is directly welfare-relevant for the voters (it is typically a policy decision). Our focus is not on how signalling leads to bad policies, but on how signalling explains reelection outcomes. Hence, it is not important to us if signalling effort has independent welfare implications. Accordingly, for simplicity we model signalling merely as a message with no direct welfare relevance (though, of course, with potential indirect welfare implications through the quality of the winner of the election).

The crucial assumption in our model is the asymmetry between the incumbent and the challenger in their ability to signal; the voters receive a signal from incumbents, but not from challengers. When the signal is interpreted as a policy outcome, it is clear why only incumbents have the ability to signal. However, since we model signalling as simply sending a message from the incumbent to the voters, another interpretation is as a form of spin, propaganda, advertising, or persuasion.[6] A large literature has documented that

[3]As will become clear it is not possible to get a causal incumbency-advantage result in a model with only two types (which is the standard in the literature): three types is the minimum necessary for the result.

[4]Our model is extremely similar to the Matthews and Mirman (1983) model of limit pricing, but that paper does not derive analogues of either our incumbency advantage or supermajority results. One advantage of having noise (which we think of as a realistic assumption) is that there exists a unique equilibrium, and therefore we do not have to invoke any of the equilibrium refinements which are usually necessary in the signalling literature.

[5]An interesting related paper is Martinez (2009). In his model politicians do not know their types, so there is no heterogeneity in effort, however effort is hump-shaped in politicians' prior reputation (which is public information), for a similar reason as effort is hump-shaped in type in our model: because politicians are trying to cross a threshold.

[6]In the political literature we find the following fitting definition of spin in Moloney (2001): "To 'spin' is to give the words describing a policy, personality or event a favourable gloss with the intention that the mass media will use them to the political advantage of the spinner and so gain public support" (page 125); "...Spin is a weak or soft form of propaganda [...] where the activity can be identified as information manipulation; where the information is more accurate than inaccurate, and where the purpose of the spin is known, to enhance the standing of the government or opposition party..."(page 128).

incumbents receive significantly more media coverage than challengers, Goldenberg and Traugott (1984) for example find that US House incumbents receive more than 60% of the combined media coverage before elections (see also Robinson (1981), Clarke and Evans (1983), Graber (1989), Kahn (1991)).[7] The widespread existence of state-owned media is an additional channel through which incumbents can receive more coverage [Besley and Prat (2006), Durante and Knight (2009)]. In fact, excess media coverage of incumbents has often been studied as a cause of incumbency advantage [Prior (2006), Snyder and Stromberg (2010)].[8] With rational expectations, extra information about the incumbent should not systematically bias voters' beliefs. Our model offers an explanation in which this casual effect can be reconciled with rational voters. Moreover it also rationalizes the phenomenon of spin itself.

A closely related paper in the persuasion literature is the recent contribution by Kamenica and Gentzkow (2011). They find a message mechanism such that even though the receiver is not fooled about the average quality of the sender, a majority of senders is perceived as more likely to be of high quality. The difference between the two papers is that, in our terminology, they allow the sender (politicians) to commit to an effort function before learning their type. Senders can therefore exogenously choose a skewed signal distribution that leads to an incumbency advantage. In contrast, we show that a skew distribution of outcomes (and thus incumbency advantage) will arise even without commitment.

Our main focus in the paper is positive, and lies in showing how signalling can lead to incumbency advantage. However, our framework can also be used to derive some normative implications, similar in spirit to Gersbach (2007; 2009; 2010), who argues for supermajority reelection rules for incumbents. In particular we show that, with incumbent signalling, voters can improve the average outcome by committing to a higher reelection threshold than that which is optimal ex post. The marginal ex ante effect of raising the threshold is to increase effort by those who were sending messages above the threshold, and decrease effort by those who were sending messages below the threshold. In the no-commitment equilibrium the types who send messages above the threshold are high quality and medium quality, while those who send messages below the threshold are low quality, thus the marginal effect of committing to a higher threshold is positive. In practice one way such a commitment could be implemented is with a constitutional amendment requiring a supermajority for an incumbent to be reelected, similar to a constitutional amendment which sets a term limit on incumbents. This mechanism is most similar to the mechanism

[7] Gordon et al. (2007) identify one way in which challengers are able to signal their type: by the choice of whether or not to enter the race.

[8] Cain et al. (1987) state that: "Incumbents win because they are better known and more favorably evaluated by any wide variety of measures. And they are better known and more favorably evaluated because, among other factors, they bombard constituents with missives containing a predominance of favorable material, maintain extensive district office operations to service their constituencies, use modern technology to target groups of constituents with particular policy interests, and vastly outspend their opponents"(p10).

in Gersbach (2009), though in his paper there are multiple equilibria, so the conclusion depends on choosing a probability distribution over equilibria. Furthermore in Gersbach (2009) the incumbent's signalling effort is welfare improving to the society. He finds that a higher vote threshold increases the average quality of reelected politicians as well as the average effort. Because the effort is welfare improving, a supermajority increases welfare. In this paper, effort is welfare neutral, and even though a supermajority always increases the average quality of reelected politicians, it also decreases the probability of reelection for good politicians. We show that this trade off is positive at the margin. This result is a contribution to the literature on optimal electoral rules (Feddersen and Pesendorfer (1998), Smart and Sturm (2006)).[9]

The remainder of the paper is organized as follows. Section 2 sets up the model, and characterizes the equilibrium. Section 3 analyzes the case of the simple majority rule. Section 4 shows that under simple majority rule an incumbency advantage exists in equilibrium. Section 5 shows that the optimal reelection rule is a supermajority rule. Section 6 proposes an illustrative calibration, and Section 7 presents further discussions of the relationship between incumbency advantage and optimal reelection thresholds, the welfare significance of signalling, issues of implementation of the supermajority rule, and possible extensions.

2 The Model

We study a game between an incumbent politician and a continuum of voters, where the voters choose between the incumbent and a challenger politician (who is a passive player in our setup). Both incumbent and challenger are defined by their talent θ which is seen as a random variable. Talent may be understood as the quality of the politician, a characteristic orthogonal to the political space, valued by every voter in the same way.[10] Examples of what might be called talent are competence and honesty. The talents of both politicians are drawn from the same distribution and are privately known only to the politicians themselves.

We assume that the distribution of talents is symmetric and has three types, i.e. $\theta \in \Theta \equiv \{\theta_L, \theta_M, \theta_H\}$, with equal distance δ between the extremes and the middle types ($\theta_H - \theta_M = \theta_M - \theta_L \equiv \delta$), and where a politician has the same probability p of being of high or low talent ($p = Pr(\theta_H) = Pr(\theta_L)$).

We have assumed symmetry in the distribution of types because we want to isolate the effect of signalling on incumbency advantage. If the underlying distribution of talent

[9]Note that a term limit on incumbents can be thought of as a special case of a majority rule, however in our model a term limit will never induce higher welfare than a simple majority, because the expected value of a chosen incumbent is bounded below by the expected value of a challenger.

[10]This concept is also called in the literature *quality* or *valence* [Ansolabehere and Snyder (2000); Carrillo and Castanheira (2002); Ashworth and Bueno de Mesquita (2009; 2008)].

was skewed, for example if the median talent was above the mean, then we would expect an incumbency advantage even without the ability to manipulate messages (because more than half of politicians would be above the expected type of the challenger, so more than half of incumbents would be re-elected). Thus the overall advantage will depend both on the skewness of the distribution of types, and the skewness in signals induced by signalling. We consider a three-type distribution because it is the minimum required to generate an incumbency advantage. Simulations with continuous distributions make us believe that the result is true more generally, but we leave derivations for future work.

The asymmetry between incumbent and challenger comes from the fact that before the election voters receive a signal of the incumbent's quality. The signal has two components, a "message" sent by the incumbent, and a noise component. The message, denoted by $\tilde{\theta}$, is in turn an additive combination of the politician's talent and his effort, $\tilde{\theta} = \theta + e$. The cost of effort is increasing and convex, denoted by $c(\cdot)$, and incurred only by the incumbent. That the message is increasing in costly effort is the key assumption of the paper. Since each incumbent's starting point is her true quality, this has the implication that lower quality incumbents must make a greater effort than higher quality ones to send the same message. One can think about this assumption as roughly capturing the idea that for an incumbent it is relatively easier to spin as a success a middling policy outcome than a policy disaster.

Voters receive the message with some noise, representing the many unobservables which contribute to political outcomes, and constrain voters' ability to infer a politician's quality. Both the incumbency advantage and the supermajority result can be derived without noise, but noise eliminates pooling and semi-separating equilibria and hence allows us to explore the comparative statics of the equilibrium. Also, noise generates a realistically continuous distribution of vote shares, which we use in our calibration. Finally, assuming that incumbents's messages are received with noise is simply realistic.

To differentiate between the information sent by the incumbent and the information received by voters, we have called *message* what the incumbent sends and *signal* what the voters receive. The signal is equal to the original message, plus noise, $s = \tilde{\theta} + \epsilon$, where ϵ is drawn from a continuous distribution with mean zero, symmetric and single peaked density distribution function $g(\cdot)$ with full support on the real line and cumulative distribution function $G(\cdot)$.

Note that all voters receive the same signal, i.e. the noise is common to all voters.[11] However, voters differ in their preferences for the incumbent. We assume that the utility of voter i given an incumbent with talent θ is given by:

$$u_i(\theta) = \theta + \eta_i \tag{1}$$

where η_i represents voter i's relative preference for the incumbent over the challenger. We

[11] We discuss in footnote 23 the general effects of heterogenous information.

assume that η_i has a continuous density $h(\cdot)$, strictly positive on $[\underline{\eta}, \overline{\eta}]$, where $\underline{\eta}$ ($\overline{\eta}$) might be minus (plus) infinity. We denote the cumulative distribution function by $H(\cdot)$ and we assume that both its mean and its median equal 0.

This model can be seen as a reduced form of a model in which, after the incumbent sends his message to the population, both incumbent and challenger (office motivated) announce their political platforms with Downsian commitment [Downs (1957)]. In any subgame perfect equilibrium of such a model, there would be convergence of platforms to the median voter's preferences, and hence the choice of effort is taken as if the voters had preferences given by (1).[12] Finally, we assume that voters support the incumbent when indifferent, though because the noise distribution is atomless, the probability of an indifference occurring is vanishingly small.

Politicians are only office-motivated. Being in office leads to a reward of π. Their only cost is the cost of effort. Thus the incumbent chooses the level of effort to maximize

$$V(\theta, e) = \pi Pr(reelection|\theta, e) - c(e)$$

The game has two decision stages. In the first stage the incumbent sends a message that the voters receive with some noise. In the second stage the voters cast their vote. The outcome of the election depends on the votes cast and the reelection rule. We will denote a reelection rule by q when the incumbent needs at least the fraction q of the votes in order to be re-elected.

In Section 3 we consider the particular case of simple majority rule for which $q = \frac{1}{2}$. Given voters' preferences a simple majority rule is equivalent to giving all power to the median voter. On the other hand, as we will discuss later in Section 5, a supermajority rule is equivalent to giving all the power to a voter who is opposed to or dislikes the incumbent. In order to be re-elected the incumbent's talent should be high enough to gain the support of this hostile voter.

Given a reelection rule q, an equilibrium is defined by an *effort rule*, $e_q : \Theta \to [0, +\infty)$ for the incumbent, and a *voting rule*, $v_q : \mathbb{R} \times [\underline{\eta}, \overline{\eta}] \to \{0, 1\}$ for the voters such that:

$$\begin{array}{ll} (i) & e_q(\theta) \in \arg\max_e \{\pi Pr_\epsilon(reelection|v_q(\cdot), \theta + e, q) - c(e)\} \\ (ii) & v_q(s, \eta_i) = 1 \quad \textit{if and only if} \quad E[\theta|s, e_q(\cdot)] + \eta_i \geq \theta_M \end{array}$$

where $Pr_\epsilon(reelection|v_q(\cdot), \theta + e, q)$ is the probability of reelection given the voting rule

[12]There is a recent literature that focuses on the interaction between the choice of effort and the choice of platform [see Ansolabehere and Snyder (2000); Aragones and Palfrey (2002); Carrillo and Castanheira (2002); Ashworth and Bueno de Mesquita (2009); Meirowitz (2008)] when there is no asymmetry between the candidates. In some of these papers there is divergence of platforms in equilibrium. We abstract from the possibility of divergence and we choose instead to work with a model that corresponds to the more standard convergence outcome because the focus is not on the interaction between valence and political competition.

$v_q(\cdot)$, the message $\tilde{\theta} = \theta + e$ and the reelection rule q, and $E[\theta|s, e_q(\cdot)]$ is the expected talent of the incumbent given that the public signal is s and using a posterior distribution of the incumbent's talent consistent with the equilibrium effort $e_q(\cdot)$.

From condition (ii), and given that all voters observe the same signal s, it is clear that whenever $v_q(s, \eta_i) = 1$ then $v_q(s, \eta_j) = 1$ for all $\eta_j \geq \eta_i$. Therefore, given a reelection rule q, the incumbent is re-elected if and only if the voter with preference $\eta_q = H^{-1}(1-q)$ supports him. The probability of reelection can be written as:

$$Pr_\epsilon(reelection|v_q(\cdot), \theta + e, q) = \int_{-\infty}^{+\infty} v_q(\theta + e + \epsilon, \eta_q) g(\epsilon) d\epsilon$$

Notice, that for some extreme reelection rules it might be the case that the outcome of the election is independent of the signal received by the voters because the preferences for or against the incumbent of the critical voter η_q outweighs any realization of the talent of the incumbent. To distinguish these uninteresting cases we denote by $\underline{q}$ ($\overline{q}$) the minimum (maximum) reelection rule such that the outcome of the reelection is not predetermined.[13]

Finally, we will say that the noise distribution $g(\cdot)$ satisfies the Monotone Likelihood Ratio Property (MLRP) if whenever $\tilde{\theta}_1 > \tilde{\theta}_2$, then $\frac{g(s-\tilde{\theta}_1)}{g(s-\tilde{\theta}_2)}$ increases in s.[14] The MLRP implies that higher signals lead to higher posterior distributions of the talent (in the sense of first-order stochastic dominance).

The following proposition gives some properties of the equilibria and states that in equilibrium the incumbent is re-elected whenever the public signal is equal to or above a certain threshold, and is not re-elected otherwise.

Proposition 1. *For any reelection rule $q \in (\underline{q}, \overline{q})$, if the cost of effort, $c(\cdot)$, is strictly convex and the distribution of noise satisfies the MLRP, then any equilibrium $e_q^*(\cdot)$ and $v_q^*(\cdot)$ satisfies the following conditions:*

(i) Voter i's best response is a threshold rule:

$$v_q(s, \eta_i) = \begin{cases} 0 & if \quad s < k_i \\ 1 & if \quad s \geq k_i \end{cases} \tag{2}$$

where k_i is determined by $E[\theta|s = k_i, e_q^*(\cdot)] + \eta_i = \theta_M$ whenever this equation has a solution, and $k_i = +\infty$ $(-\infty)$ if $E[\theta|s, e_q^*(\cdot)] + \eta_i < (>)\ \theta_M$ for all $s \in \mathbb{R}$. Moreover, k_i is decreasing in the preference parameter η_i.

(ii) The incumbent is re-elected if and only if the public signal is above a threshold k_q, where k_q is given by:

$$E[\theta|s = k_q, e_q^*(\cdot)] = \theta_M - H^{-1}(1-q) \tag{3}$$

[13] More precisely, $\underline{q} = 1 - H(\delta)$ and $\overline{q} = 1 - H(-\delta)$.

[14] This definition corresponds to the special case of the MLRP defined by Milgrom (1981) when the signal structure is additive.

(iii) The incumbent's optimal effort solves:

$$\begin{aligned} \pi g(\theta + e_q^*(\theta) - k_q) &= c'(e_q^*(\theta)) \\ \pi g'(\theta + e_q^*(\theta) - k_q) - c''(e_q^*(\theta)) &< 0 \end{aligned} \tag{4}$$

Moreover, the incumbent's message is increasing in his type (i.e. $\tilde{\theta}_L < \tilde{\theta}_M < \tilde{\theta}_H$*)*

Part (i) of Proposition 1 states that in equilibrium the voters follow a threshold rule such that they support the incumbent if and only if the signal received is above a threshold that is decreasing in their preferences for the incumbent. At the threshold k_i, voter i is indifferent between supporting the incumbent and appointing a new politician.

Part (ii) states that for reelection, the signal received must be above the threshold k_q that leaves indifferent the critical voter with preference $\eta_q = H^{-1}(1-q)$.

Part (iii) states the first and second order conditions for the optimal effort given that the expected payoff of the incumbent can now be written as:

$$V(\theta, e, q) = \pi G(\theta + e - k_q) - c(e)$$

In particular the optimal level of effort is such that the marginal increase in the probability of reelection should equal the marginal cost of effort.

It is important to note that the incumbent's message is increasing in his type. Intuitively, all incumbents (independent of their type) have the same reward from re-election, however it is more costly for low talented incumbents to send high messages. If an incumbent finds worth to exert some effort to reach the reward, then those more talented will strictly benefit to match that message. The monotonicity of the message together with the MLRP assumption implies that the expected talent given the realization of a signal is increasing on the signal, and will be widely used in the rest of the paper.

Finally, to guarantee that the local first and second order condition are sufficient for a global optimum we assume throughout the paper the following condition:

$$\inf_e c''(e) > \pi \sup_\epsilon g'(\epsilon) \tag{5}$$

Condition (5) requires the cost function to be sufficiently convex, so that the marginal cost cuts only once the marginal benefit.

The proof of Proposition 1 is similar to that of Theorem 1 in Matthews and Mirman (1983) regarding limit pricing and therefore is relegated to the Appendix.[15]

[15]The setup in Matthews and Mirman (1983) is close to ours: a monopoly wants to deter the entrant of a possible challenger, and they do so by lowering their price, to signal lower profitability in the

3 Simple Majority Rule

As a benchmark consider the simple majority rule $q = \frac{1}{2}$. Notice that given the assumptions on the voters' preferences for the incumbent, equation (3) becomes:[16]

$$E[\theta | s = k^*] = \theta_M \tag{6}$$

where k^* denotes the equilibrium threshold in the simple majority case. In other words, the simple majority rule is equivalent to giving all the power to the median voter, the voter that is ex-ante (before receiving the signal) indifferent between the incumbent and the challenger. The incumbent will be re-elected if and only if this voter believes him to have a higher than average talent.

The equilibrium for the simple majority rule has the following properties:

Proposition 2. *With a simple majority rule, the equilibrium is unique. The effort levels satisfy $e_M > e_L = e_H \equiv e^*$ with $e^* = c'^{-1}(\pi g(\theta_H - \theta_M))$ and the threshold signal is given by $k^* = \theta_M + e^*$.*

Proof For clarity we omit the reference to the electoral rule on the equilibrium variables. Given the talent distribution, upon receiving a signal $s = k^*$, equation (6) becomes:

$$\frac{\sum_j \theta_j g(k^* - \tilde{\theta}_j) Pr(\theta_j)}{\sum_j g(k^* - \tilde{\theta}_j) Pr(\theta_j)} = \theta_M$$

and given $Pr(\theta_H) = Pr(\theta_L)$ and $\theta_H - \theta_M = \theta_M - \theta_L$, it simplifies to:

$$g(k^* - \tilde{\theta}_H) = g(k^* - \tilde{\theta}_L) \tag{7}$$

The first order conditions for the equilibrium effort (4) together with equation (7) imply that $e_H = e_L$. Denote by e^* this effort level.

Then given the symmetry of the noise distribution, equation (7) implies that the equilibrium threshold will be exactly half-way between the signals sent by the high and low type incumbents:

$$k^* = \frac{\tilde{\theta}_H + \tilde{\theta}_L}{2} = \theta_M + e^* \tag{8}$$

To see that $e_M > e^*$ notice that, from the single-peakedness and symmetry of g:

$$\pi g(k^* - \theta_M - e^*) > \pi g(k^* - \theta_L - e^*) = c'(e^*)$$

market. Analogously, a politician exerts effort to signal their type. Note that Proposition 1 is true for an arbitrary distribution of types, not just for our three-type distribution. In particular, it would be true for a continuous distribution of types.

[16] For clarity we suppress reference to the effort function $e^*_{1/2}(\cdot)$.

that is, the marginal benefit for an incumbent with type θ_M of exerting effort e^* outweighs the marginal cost of exerting this level of effort. Therefore, $e_M > e^*$.

Finally, replacing $k^* = \theta_M + e^*$ into the first order conditions for e^* given by equation (4), we obtain the equilibrium level e^*:

$$c'(e^*) = \pi g(\delta)$$

where $\delta \equiv \theta_H - \theta_M = \theta_M - \theta_L$ represents the dispersion of the talent distribution[17]. $\square$

The proof of Proposition 2 can be visualized in Figure 1. Both the talents and the messages can be read on the horizontal axis. The upward sloping lines represent the marginal costs of effort for each type. Consider the messages sent by the low-quality and high-quality incumbents, $\tilde{\theta}_L$ and $\tilde{\theta}_H$. By Proposition 1, $\tilde{\theta}_L < \tilde{\theta}_H$. Since the median voter is indifferent between middle types and the challenger, she will be indifferent between the incumbent and challenger only when the signal is equally likely to be from either a high or a low type. Therefore using the symmetry of the noise distribution she will set the threshold of acceptance k^* at the midpoint between $\tilde{\theta}_L$ and $\tilde{\theta}_H$. Given k^*, the marginal benefit of effort at $\tilde{\theta}$ is $\pi g(\tilde{\theta} - k^*)$, i.e. the marginal increase in the probability of being reelected, which is also depicted in Figure 1. By the symmetry of g, the marginal benefit is symmetric around k^*, and since k^* is the middle point between $\tilde{\theta}_L$ and $\tilde{\theta}_H$, the marginal benefits of effort at both points are equal. In equilibrium marginal benefits equal marginal costs, and hence the equilibrium efforts exerted by the low-quality and high-quality incumbents must be the same $e^*_L = e^*_H = e^*$. Finally, returning to the middle types, if they exerted the same effort e^* they would face a higher marginal benefit of effort, due to the single-peakedness of the of the marginal benefit function. Thus middle types increase their effort till their marginal cost equates their marginal benefit, at $\tilde{\theta}_M$. Note that although the message $\tilde{\theta}$ is strictly increasing in an incumbent's type, the effort is non-monotonic in types.

The effort level e^* is increasing in π and decreasing in the dispersion of the incumbent's talent δ. These results are very intuitive, a direct change in the marginal benefit or cost changes the effort level accordingly. Moreover, if the distance between incumbents increases then it is more difficult to fool the voters by exerting effort and therefore the marginal benefit of effort goes down and they exert less effort.

The medium type's effort level e^M is also decreasing in the dispersion δ, because a lower e^* corresponds to a lower k, and therefore a lower incentive to exert effort. The medium type's effort is also increasing in π, both because of the direct effect (a higher incentive to exert effort), and through its indirect effect of raising e^*, and therefore k.

Assuming that the noise is normally distributed with variance σ^2_ϵ and mean zero, we

[17] The distance between the talents of the incumbents is a measure of the dispersion of the distribution. In fact the variance of the talents is given by: $Var(\theta) = 2p(\theta_H - \theta_M)^2 = 2p\delta^2$

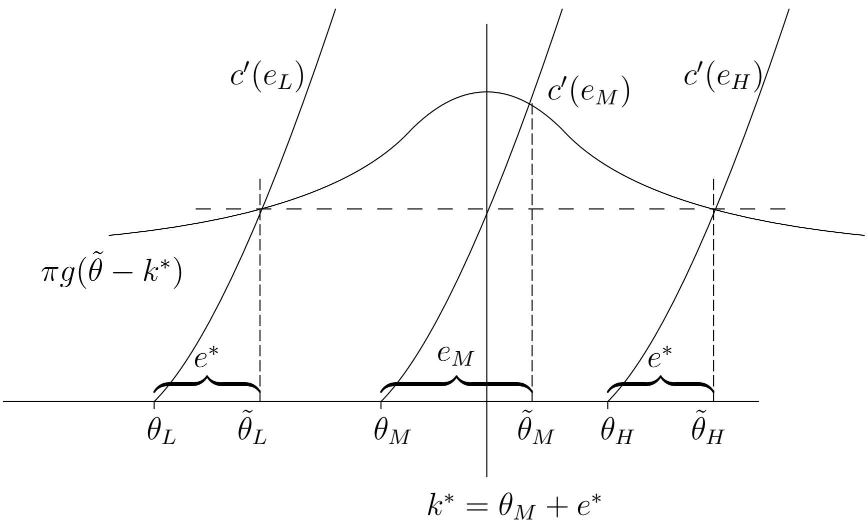

Figure 1: Equilibrium

can further study how the equilibrium effort level changes with the variance of the noise. The change in the equilibrium effort with respect to the variance of the noise depends on the relative size of the variance of the noise and the square of the dispersion of the incumbents:

$$\partial e^*/\partial \sigma_\epsilon^2 < 0 \quad \text{if and only if} \quad \sigma_\epsilon^2 > \delta^2 \tag{9}$$

To understand this result consider the following two extreme scenarios. Suppose that the signal is *extremely noisy*, then voters do not infer much from the signal and incumbents exert very little effort. If the variance of the signal decreases making the signal more informative, then reelection will be more responsive to the signal received and incumbents will exert more effort. On the other hand, if the signal is *very precise*, incumbents are not going to be able to fool the voters and exert little effort. Condition (9) says that whether we consider the signal *extremely noisy* or *very precise* depends on the relative variances of the two distributions.

4 Incumbency Advantage

One interesting feature of the equilibrium is that the incumbents with middle talent are the ones that exert higher effort. The reason is that the equilibrium threshold is closer to their types and hence they have greater incentive to exert effort. This extra effort from the incumbents with middle talent implies that the distribution of the messages, signals, and ultimately of expected types, will be negatively skewed (the median is above the mean) leading to our result of an incumbency advantage.

In the introduction we have characterized a causal incumbency advantage as a situation in which, on average, incumbents have a better-than-50% chance of reelection against a challenger with similar intrinsic characteristics. Because the quality distribution of incumbents running for reelection is identical to the quality distribution of challengers, in our model this is equivalent to saying that on average incumbents are reelected more than 50% of the time, and this will be our definition of incumbency advantage.

Proposition 3. *Under simple majority there is an incumbency advantage.*

Proof From Proposition 2, $e_M > e^*$. The probability of reelection for an incumbent with talent θ_j that sends message $\tilde{\theta}_j$ is then:

$$Pr(reelection \mid \tilde{\theta}_j) = Pr(\tilde{\theta}_j + \epsilon > k^*) = 1 - G(k^* - \theta_j - e_j)$$

The unconditional probability of reelection is therefore:

$$\begin{aligned} Pr(s \geq k^*) &= p(1 - G(k^* - \tilde{\theta}_H)) + p(1 - G(k^* - \tilde{\theta}_L)) + (1-2p)(1 - G(k^* - \tilde{\theta}_M)) \\ &= p + (1-2p)(1 - G(k^* - \tilde{\theta}_M)) \\ &> \tfrac{1}{2} \end{aligned} \tag{10}$$

Where the second equality follows because $G(k^* - \tilde{\theta}_H) = 1 - G(k^* - \tilde{\theta}_L)$ and the inequality because $e_M > e^*$ so $\tilde{\theta}_M > k^*$. □

Intuitively, when the median voter chooses whether to reappoint the incumbent or not, she compares her updated belief about the talent of the incumbent with the expected talent of the challenger. In doing so she can ignore middle type incumbents because they have just average talent, and hence taking into account the equilibrium messages of the incumbents, the threshold signal would be just the middle point between the messages sent by the low and the high signals. But given that the incumbents with middle talent exert more effort than the others, the message $\tilde{\theta}_M$ will exceed the threshold and therefore they will be re-elected in more than half of their attempts.

It is important to note that in this model there is an incumbency advantage *on average*. In particular, high-quality incumbents will win a majority of elections, and low-quality incumbents will lose a majority of elections (as the voters have no information on the quality of the challenger, so the reelection depends exclusively on the expected quality of the incumbent). Because of the symmetry in the distribution of types and the fact that high- and low-quality incumbents exert the same effort, the overall combined reelection rates of low- and high-quality incumbents is exactly 50%. Hence, whether or not there is on average an incumbency advantage depends on whether or not middle-type incumbents are re-elected more than 50% of the time. Since middle-type incumbents send a message that is above the voters' reelection threshold, a majority of them will indeed be reelected. This is fully consistent with the empirical literature on causal incumbency advantage. That literature captures the incumbency advantage among candidates of similar quality, but, as quality is unobservable, it cannot estimate separate incumbency advantages for different levels of quality. Instead, it can only estimate the average advantage across quality pairs. Hence, it is perfectly possible that at low quality levels there is an incumbency disadvantage, as in our model.

Relatedly, it is also important to note that just because signalling leads to an incumbency advantage does not imply that signalling is socially harmful. Consider what would happen in our model if signalling were not possible (equivalently, if effort was infinitely costly). The reelection threshold would again be halfway between the low and high messages, but there would be no incumbency advantage, because the middle type would not exert any effort. However voter welfare would be the same as in the case with signalling, because although fewer middle types would be reelected, the welfare value of middle types is neutral anyway.[18]

5 Supermajority

In this section we consider the social planner's problem of maximizing the total welfare of the voters by choosing a reelection rule (we treat the welfare of the incumbent politician as negligible when computing the social welfare). We prove that the simple majority rule is suboptimal and that the welfare maximizing rule must be a supermajority rule ($q > \frac{1}{2}$).[19]

We proceed in two steps. First, we show that under simple majority voters would be better off if they could commit to a higher threshold to re-elect the incumbent. This commitment is not credible because ex post it is efficient to re-elect the incumbent if the updated beliefs indicate that he is above average (i.e., if the median voter would prefer him). We then propose a way to implement this commitment by setting a supermajority rule that takes decision power from the median voter and gives it to a voter with a partisan position somewhat against the incumbent.

Proposition 4. *Under simple majority, the welfare maximizing reelection threshold is above the equilibrium threshold k^*.*

Proof Given a threshold k, the expected welfare can be expressed as the value of the outside option (the expected value of a challenger, θ_M), plus the expected change in value

[18]The neutrality of signalling for welfare purposes may be specific to our three-type models. For general distributions of types, under simple majority the kind of signalling we model could be welfare improving or welfare reducing. The point we are making here is that it is not necessarily welfare reducing. What we believe is always true is that irrespective of the welfare consequences of signalling under simple majority, welfare can always be improved by increasing the reelection threshold, which is the subject of the next section.

[19]This seems to contradict the literature on biased contests [Meyer (1991; 1992)] in which favoring early success by increasing the probability of success in later contests might be optimal. However, in the setup of those papers, either no action (effort) was required from the contesters, or all the contester had the same known ability (no types), and hence the incentive effect that arise from the supermajority rule in the present paper cannot be present there.

from retaining the incumbent:[20]

$$\begin{aligned} EW &= \theta_M + pPr(\tilde{\theta}_H + \epsilon \geq k)(\theta_H - \theta_M) + pPr(\tilde{\theta}_L + \epsilon \geq k)(\theta_L - \theta_M) \\ &= \theta_M + p\delta(G(\tilde{\theta}_H - k) - G(\tilde{\theta}_L - k)) \end{aligned} \tag{11}$$

The optimal threshold is then determined by the first order condition:

$$\frac{\partial EW}{\partial k} = p\delta\left(g(\tilde{\theta}_H - k)(\frac{\partial e_H}{\partial k} - 1) - g(\tilde{\theta}_L - k)(\frac{\partial e_L}{\partial k} - 1)\right) = 0 \tag{12}$$

At the equilibrium threshold, $g(\tilde{\theta}_H - k^*) = g(\tilde{\theta}_L - k^*)$, therefore if we evaluate the derivative (12) at k^*, the direct effect on welfare of a change in the threshold is zero. However, the change in the threshold also affects the choice of effort. Recall that the optimal level of effort given a threshold k satisfies the following first and second order conditions:

$$\begin{aligned} \pi g(\theta_j + e_j - k) &= c'(e_j) \\ \pi g'(\theta_j + e_j - k) - c''(e_j) &< 0 \end{aligned} \tag{13}$$

In particular, totally differentiating the first order condition with respect to k and rearranging:

$$\frac{\partial e_j}{\partial k} = \frac{\pi g'(\theta_j + e_j - k)}{\pi g'(\theta_j + e_j - k) - c''(e_j)} \tag{14}$$

and using the second order condition and the fact that $g'(\theta_L + e_L - k^*) > 0 > g'(\theta_H + e_H - k^*)$ we have that:

$$\left.\frac{\partial e_H}{\partial k}\right|_{k=k^*} > 0 \quad and \quad \left.\frac{\partial e_L}{\partial k}\right|_{k=k^*} < 0$$

Hence, plugging this into (12), the indirect effect on welfare of a raise in the threshold is positive. Increasing the threshold causes θ_H to exert more effort[21] while θ_L will reduce his effort, leading to more separation between the incumbents' signals and as a result an increase in welfare.

We have shown that welfare is improved by marginally increasing the threshold from its Nash equilibrium level. However this is not sufficient to show that a threshold higher than the Nash equilibrium threshold is optimal, because the welfare function may not be single-peaked. We therefore demonstrate below that for any threshold $k < k^*$ the correspondent welfare is strictly lower than the welfare at the equilibrium threshold k^*. To see that, first notice that given θ_j, the optimal effort level e_j defined by equation (13) is a single-peaked function of the threshold k. For a given θ_j, the effort $e_j(\cdot)$ is increasing for $k < \theta_j + c'^{-1}(\pi g(0))$ and decreasing otherwise. Moreover, given equation (13), we have

[20] The partisan preferences (η_i) disappear from this expression, because of their zero mean.

[21] A marginally higher threshold also leads the middle θ_M to exert more effort. To see this observe that $\theta_M + e_M - k^* = e_M - e^* > 0$ and hence $g'(\theta_M + e_M - k^*) < 0$ and $\frac{\partial e_M}{\partial k}|_{k=k^*} > 0$.

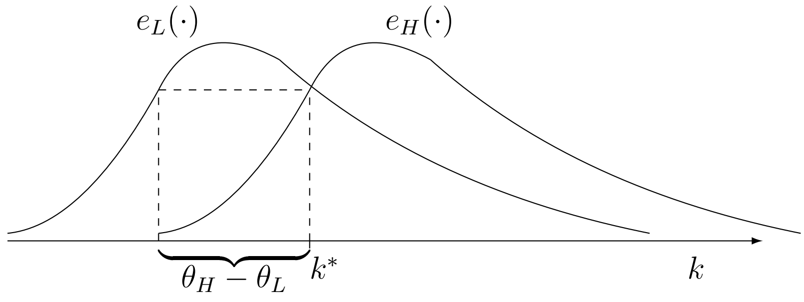

Figure 2: Effort functions

the following identity:

$$e_L(k - (\theta_H - \theta_L)) \equiv e_H(k)$$

so the optimal effort function of the low type is a horizontal shift to the left of the effort of the high type (see Figure 2).

At the equilibrium threshold, $e_L(k^*) = e_H(k^*) \equiv e^*$ so $e_L(k^*) = e_L(k^* - (\theta_H - \theta_L))$ which implies that k^* is on the downward-sloping part of curve $e_L(\cdot)$ and on the upward-sloping part of $e_H(\cdot)$. A representation of the effort functions can be seen in Figure 2.

Consider $k < k^*$, then $e_L(k) > e_H(k)$ and hence the distance between the high and low messages under threshold k is smaller than under threshold k^*:

$$\tilde{\theta}_H(k) - \tilde{\theta}_L(k) < \tilde{\theta}_H^* - \tilde{\theta}_L^* \tag{15}$$

Notice that by the symmetry of the noise distribution, the following two remarks are satisfied:

R1: Whenever two points are at a fixed distance h, $G(x) - G(x-h)$ is maximized at $x = \frac{h}{2}$, that is, when the two points are equidistant to the mean.[22]

R2: Given two points equidistant to the mean, the difference in the cumulative distribution is increasing in the distance between the two points:

$$\frac{\partial}{\partial h}\left[G(\frac{h}{2}) - G(-\frac{h}{2})\right] = \frac{1}{2}(g(\frac{h}{2}) + g(-\frac{h}{2})) > 0$$

We can now conclude that for any threshold $k < k^*$ the welfare under threshold k is

[22]To see this consider the first order condition with respect to x: $g(x) - g(x-h) = 0$, and by the symmetry of $g(\cdot)$, this implies $x = -(x-h)$ or $x = \frac{h}{2}$.

lower than under the equilibrium threshold k^*:

$$\begin{aligned}
EW(k) &= \theta_M + p\delta(G(\tilde{\theta}_H(k) - k) - G(\tilde{\theta}_L(k) - k)) \\
&\leq \theta_M + p\delta(G(\tfrac{\tilde{\theta}_H(k) - \tilde{\theta}_L(k)}{2}) - G(-\tfrac{\tilde{\theta}_H(k) - \tilde{\theta}_L(k)}{2})) \\
&\leq \theta_M + p\delta(G(\tfrac{\tilde{\theta}^*_H - \tilde{\theta}^*_L}{2}) - G(-\tfrac{\tilde{\theta}^*_H - \tilde{\theta}^*_L}{2})) \\
&= EW(k^*)
\end{aligned}$$

where the first inequality follows from R1 and the second from R2 and (15). □

Proposition 4 implies that the voters would be better off if they could commit to re-elect incumbents that have expected talent above a level which is strictly higher than the ex-ante average talent. An increase in the threshold will cause high types to exert more effort and low types to exert less effort. For both types their efforts will not offset the increase in the threshold, so both will be re-elected with a lower probability. But it is the larger fall in the probability of low-type reelection that increases welfare.

This higher threshold is not optimal ex post, because it asks the voters to not re-elect some politicians with expected talent strictly greater than the expected talent of the challenger. It is not clear that individual voters have access to credible commitment devices, allowing them to implement the higher threshold. However committing to a higher threshold has a natural interpretation with respect to the electorate as a whole: a constitutional rule such that incumbents will only be allowed a second term if they exceed some threshold of the vote share strictly greater than one half, i.e. a supermajority rule.

If all voters are identical then this rule, of course, has no effect. However, if the voters differ in their preferences for the incumbent, in the way we have assumed, then a supermajority rule transfers the decision power from the median voter to a voter that is ideologically opposed to the incumbent.[23] Therefore a supermajority rule acts in effect as a commitment device that sets a higher threshold of talent for reelection.

Proposition 5. *The welfare maximizing reelection rule is a supermajority rule ($q_W > \frac{1}{2}$).*

Proof Given a threshold k, there is a reelection rule that implements that threshold in equilibrium. Denote by $e_k(\cdot)$ the optimal effort the incumbent exerts if he faces threshold k,[24] as a function of his type. We define $q(k)$ as follows:

$$q(k) = 1 - H(\theta_M - E[\theta | s = k, e_k(\cdot)]) \tag{16}$$

[23] Another source of voter heterogeneity may be differential information. However if agents are rational, and there is common knowledge of rationality, then it is difficult to argue that the heterogeneous information will not be efficiently aggregated. Information can be indirectly passed through, for example, opinion polls. If a voter compares her own private signal with the aggregated signals of 1000 people in an opinion poll, then the latter would seem to swamp the former. Also voters should vote using the expectations conditional on being decisive; this force will generally make a supermajority rule less effective [see Feddersen and Pesendorfer (1998)].

[24] The effort function $e_k(\cdot)$ solves equation (13).

Clearly, setting the reelection rule $q = q(k)$ leads to the equilibrium effort $e^*_{q(k)}(\cdot) \equiv e_k(\cdot)$ and to the equilibrium threshold $k_{q(k)} = k$. To prove Proposition 5 it would be sufficient to prove that $q(k)$ is increasing in k. However this need not be true everywhere. As the threshold gets past a certain point both high and low types will react to an increase in the threshold by lowering their levels of effort (see Figure 2), thus an increase in the threshold could correspond to a lower expected quality from a signal sent at the threshold.

To prove the result we proceed in two steps. First we note that, by Proposition 2, the equation $q(k) = \frac{1}{2}$ has a unique solution at k^*, the equilibrium threshold of a simple majority case. Then we show that $q(\cdot)$ is strictly increasing at k^*, the equilibrium threshold of the simple majority case. Since $q(\cdot)$ is continuous, and only cuts the line $q(k) = \frac{1}{2}$ once, this implies that for any $k > k^*$, $q(k) > q(k^*) = \frac{1}{2}$.

Formally, $q(k) = \frac{1}{2}$ if and only if $E[\theta|s = k, e_k(\cdot)] = \theta_M$. By equation (6), k^* satisfies $E[\theta|s = k^*, e_{k^*}(\cdot)] = \theta_M$. To see that k^* is the unique solution to this equation notice that if $E[\theta|s = k, e_k(\cdot)] = \theta_M$, it has to be the case that $\tilde{\theta}_L(k)$ and $\tilde{\theta}_H(k)$ are equidistant to the threshold k. This implies that $e_k(\theta_L) = e_k(\theta_H) = k - \theta_M$. Substituting this in the first order conditions leads to $e_k(\theta_L) = e^* = c'^{-1}(\pi(\theta_H - \theta_M))$ and $k = k^*$.

We now show that $q(k)$ is increasing at k^*, or equivalently, that $E[\theta|s = k, e_k(\cdot)]$ is increasing at k^*:

$$\left.\frac{\partial E[\theta|s = k, e_k(\cdot)]}{\partial k}\right|_{k=k^*} = \frac{\partial}{\partial k}\left.\left[\frac{(\theta_H - \theta_M)p[g(\tilde{\theta}_H - k) - g(\tilde{\theta}_L - k)]}{p(g(\tilde{\theta}_H - k) + g(\tilde{\theta}_L - k)) + (1 - 2p)g(\tilde{\theta}_M - k)}\right]\right|_{k=k^*}$$

Denoting by D the denominator of this fraction:

$$\begin{aligned}\left.\frac{\partial E[\theta|s=k,e_k(\cdot)]}{\partial k}\right|_{k=k^*} &= \frac{(\theta_H - \theta_M)p}{D}\left[g'(\tilde{\theta}_H - k^*)(\tfrac{\partial e_H(k^*)}{\partial k} - 1) - g'(\tilde{\theta}_L - k^*)(\tfrac{\partial e_L(k^*)}{\partial k} - 1)\right] \\ &> 0\end{aligned}$$

where the inequality follows because $D > 0$, $g'(\tilde{\theta}_H - k^*) = -g'(\tilde{\theta}_L - k^*) < 0$ by the equilibrium condition (11) and $\frac{\partial e_j(k^*)}{\partial k} < 1$ for $j \in \{H, L\}$ by equation (14).

Therefore, denoting by k_W the welfare maximizing threshold defined by equation (12), $k_W > k^*$ by Proposition 4 and therefore the optimal reelection rule $q(k_W) > \frac{1}{2}$ is a supermajority rule. □

6 Numerical Illustration

In this section we present a simple numerical exercise to illustrate the potential magnitudes involved in our model. The exercise has two modest goals. First, to show that a set of parameters which seem intuitively reasonable (to the authors at least) can reproduce incumbency effects of the right order of magnitude. Second, to show that the implications for an optimal supermajority rule and its welfare effects are also of an intuitive magnitude.

We assume that the noise and preference distributions are normal. We also assume a quadratic cost of effort function, $c(e) = \frac{c}{2}e^2$, and without loss of generality we set $\theta_M = 0$ and $\pi = 1$. The model has then five free parameters: (1) the variance of the noise distribution σ_ϵ^2, (2) the variance of the voters' preferences σ_η^2, (3) the dispersion of the talent distribution $\delta = \theta^H - \theta^M$, (4) the probability of the high and low types p, and (5) the parameter c in the cost-of-effort function.[25]

Numerical experimentation, documented below, shows that the optimal reelection threshold is fairly insensitive to three of these five parameters: namely c, σ_ϵ^2, and δ. We thus fix these parameters at the arbitrary values of 0.25, 1, and 1.5. We then calibrate σ_η^2 and p (to both of which the optimal supermajority rule is quite sensitive) by targeting the causal incumbency advantage numbers reported in Lee (2008). That paper uses a regression discontinuity analysis on U.S. Congressional elections, and finds that the difference in the probability of winning an election between a marginal winner and a marginal loser (i.e., a winner or loser of the previous election) is 35%, and that the average difference in vote shares is of 7%.[26] The model's formulas for probability of reelection and vote share are reported in the appendix. Solving these formulas to match Lee's estimates yields the probabilities of the low and high type to be $p = 0.165$ and a standard deviation for voters' ideological preferences of $\sigma_\eta = 0.6$. With these parameters, using equations (12) and (16), the optimal supermajority rule is $q_W = 57\%$. This supermajority rule leads to a welfare increase of 3.35% relative to simple majority.

In Figure 3 we plot the optimal supermajority rule as a function of each of the parameters, holding all remaining parameters fixed at the benchmark level. As can be seen the optimal reelection threshold is fairly insensitive to c, σ_ϵ^2, and δ, but not to p and σ_η.[27] Figure 4 shows voters' welfare as a function of the reelection rule: a supermajority dominates simple majority for all supermajorities less than 61%.

7 Discussion and Conclusions

This paper shows that in a model with noisy signalling and a threshold rule (which seems a natural treatment of elections) the distribution of signals will be skew, such that the expected type of the median signal will be above average, meaning that more than half of incumbents will have expected type greater than the challenger, thus generating an incumbency advantage.

[25] The sufficient condition (5) is translated in the following restriction for the parameters:

$$c \geq \frac{1}{\sigma_\epsilon^2\sqrt{2\pi}}e^{-\frac{1}{2}}$$

[26] These numbers correspond to the party rather than the candidate incumbency advantage and average vote share advantage. The problem with the establishment of a candidate incumbency advantage is that there is an endogenous attrition of candidates that distorts the results.

[27] We restrict the support of c and σ_ϵ to those values satisfying the sufficient condition 5.

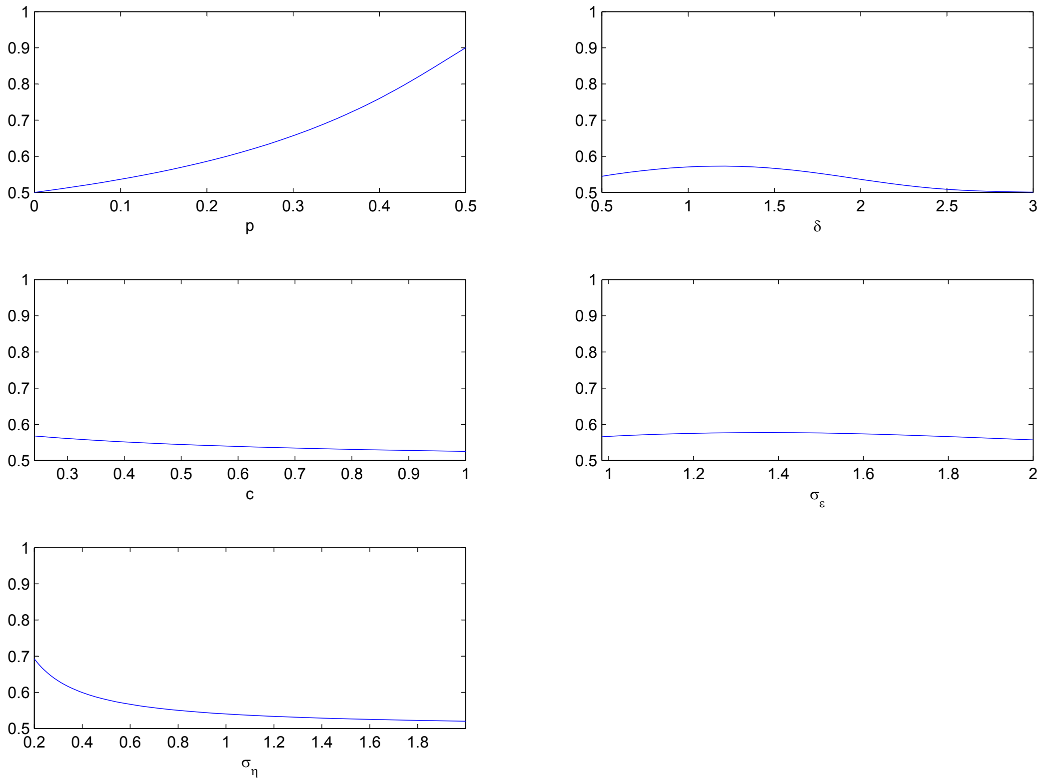

Figure 3: Optimal supermajority rule for changes in the parameters. Default parameters: $c = \frac{1}{4}$, $\delta = 1.5$, $p = 0.165$, $\sigma_\epsilon = 1$, $\sigma_\eta = 0.6$

We have derived these results using a three-type model. A natural question is whether our results would extend to more general distributions of types. As already mentioned, extensive numerical calculations suggest that both the positive and the normative predictions of the model are robust to a wide range of discrete and continuous distributions of types, though a formal proof has eluded us thus far. Our conjecture, therefore, is that incumbency advantage and the optimality of supermajority reelection thresholds are generic features of models with noisy signalling by incumbents (as long as the signalling has no direct welfare costs, of course. If there are direct costs the supermajority results will depend on the relative benefits of improved screening and the cost of the signalling action).[28]

We reiterate that in our model incumbency advantage is not a "problem", nor is the supermajority reelection threshold a solution to a pathology. However incumbency advantage is an indication that incumbents engage in signalling, and voters have then an incentive to take advantage of such signalling by choosing the reelection threshold ap-

[28] A special case is when there are only two types ($p = 0.5$ in our model). Then there is no incumbency advantage. It is still optimal to have a supermajority reelection rule, though.

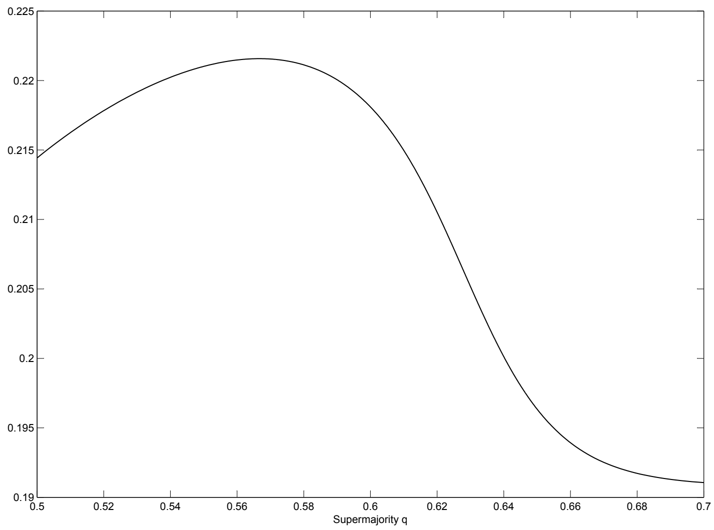

Figure 4: Welfare for different supermajority rules. ($c = \frac{1}{4}$, $\delta = 1.5$, $p = 0.165$, $\sigma_\epsilon = 1$, $\sigma_\eta = 0.6$.)

propriately. Note, too that the aim of the supermajority reelection threshold is not to eliminate the incumbency advantage. In our calibration, a small incumbency advantage remains even at the optimal threshold.[29]

As mentioned in the Introduction, our model can be seen as an extension of the noisy signalling model of Matthews and Mirman (1983) which applies to limit pricing. Besides limit pricing our approach may be fruitful in a number of other contexts in which thresholds are observed. A natural analogue to elections is a hiring decision: internal candidates may face an advantage simply due to their ability to signal [Chan (1996)], and firms may therefore find it optimal to handicap internal candidates, to improve the separation of types. More generally the results could be applied to competitions which award a prize for demonstrating an ability which exceeds some threshold, such as acceptance into a program conditional on the score on a standardized test. Without commitment, more applicants will pass the test than would pass under full information; and with commitment, the administrators of the competition have an incentive to announce a higher threshold.

[29]Nor is it true, however, that at the optimal reelection threshold there will always be an incumbency advantage. For example, when the probability of the middle type is zero, a supermajority threshold lowers the probability of reelection for both types, thus lowering the average probability of reelection below $\frac{1}{2}$.

References

Acemoglu, D., Egorov, G. and Sonin, K. (2011). A Political Theory of Populisim. MIT working paper 11–21.

Ansolabehere, S. and Snyder, J. M. (2000). Valence politics and equilibrium in spatial election models, *Public Choice* **103**: 327–336.

Ansolabehere, S., Snyder, J. M. and Stewart, C. (2000). Old Voters, New Voters, and the Personal Vote: Using Redistricting to Measure the Incumbency Advantage, *American Journal of Political Science* **44**: 17–34.

Aragones, E. and Palfrey, T. R. (2002). Mixed Equilibrium in a Downsian Model with a Favoured Candidate, *Journal of Economic Theory* **103**: 131–161.

Ashworth, S. and Bueno de Mesquita, E. (2008). Electoral Selection, Strategic Challenger Entry, and the Incumbency Advantage, *The Journal of Politics* **70**: 1006–1025.

Ashworth, S. and Bueno de Mesquita, E. (2009). Elections with platform and valence competition., *Games and Economic Behavior* **67**(1): 191–216.

Besley, T. and Prat, A. (2006). Handcuffs for the Grbbing Hand?: Media Capture and Government Accountability, *American Economic Review* **93**(3): 720–736.

Cain, B., Ferejohn, J. and Fiorina, M. (1987). *The Personal Vote: Constituency Service and Electoral Independence*, Harvard University Press.

Canes-Wrone, B., Herron, M. C. and Shotts, K. W. (2001). Leadership and Pandering: A Theory of Executive Policymaking, *American Journal of Political Science* **45**(3): 532–550.

Carrillo, J. D. and Castanheira, M. (2002). Platform Divergence, Political Efficiency and the Median Voter Theorem. CEPR Discussion Paper No. 3180.

Chan, W. (1996). External recruitment versus internal promotion, *Journal of Labor Economics* **14**(4): 555–570.

Clarke, P. and Evans, S. H. (1983). *Covering Campaigns: Journalism in Congressional Elections*, Stanford, CA: Stanford University Press.

Dick, A. R. and Lott, J. R. (1993). Reconciling voters' behavior with legislative term limits, *Journal of Public Economics* **50**(1): 1–14.

Downs, A. (1957). *An Economic Theory of Democracy*, New York: Harper and Row.

Durante, R. and Knight, B. (2009). Partisan Control, Media Bias and Viewer Responses: Evidence from Berlusconi's Italy. NBER Working Paper 14762.

Feddersen, T. and Pesendorfer, W. (1998). Convicting the Inocent: The Inferiority of Unanimous Jury Verdicts under Strategic Voting, *American Political Science Review* **92**(1): 23–35.

Gelman, A. and King, G. (1990). Estimating incumbency advantage without bias, *American Journal of Political Science* **34**(1): 1142–1164.

Gersbach, H. (2007). Vote-Share Contracts and Democracy. CEPR Discussion Paper DP6497.

Gersbach, H. (2009). Higher Vote Thresholds for Incumbents, Effort and Selection. CEPR Working Paper.

Gersbach, H. (2010). On Higher Hurdles for Incumbents, *Economics Bulletin* **30**(1).

Goldenberg, E. and Traugott, M. (1984). Campaigning for Congress, *Washington, DC: Congressional Quarterly Press* .

Gordon, S. C., Huber, G. A. and Landa, D. (2007). Challenger entry and voter learning, *American Political Science Review* **101**(2): 303.

Graber, D. (1989). Mass Media and American Politics, *Washington, DC: Congressional Quarterly Press* .

Kahn, K. F. (1991). Senate Elections in the News: Examining Campaign Coverage, *Legislative Studies Quarterly* **16**: 349–374.

Kamenica, E. and Gentzkow, M. (2011). Bayesian Persuasion, *American Economic Review* **101**(6): 2590–2615.

Lee, D. S. (2008). Randomized experiments from non-random selection in U.S. House elections, *Journal of Econometrics* **142**: 675–697.

Levitt, S. D. and Wolfram, C. D. (1997). Decomposing the Sources of Incumbency Advantage in the U.S. House, *Legislative Studies Quarterly* **22**: 45–60.

Martinez, L. (2009). A theory of political cycles, *Journal of Economic Theory* **144**(3): 1166–1186.

Maskin, E. and Tirole, J. (2004). The Politician and the Judge: Accountability in Government, *The American Economic Review* **94**(4): 1034–1054.

Matthews, S. A. and Mirman, L. J. (1983). Equilibrium Limit Pricing: The Effects of Private Information and Stochastic Demand, *Econometrica* **51**(4): 981–996.

Meirowitz, A. (2008). Electoral Contests, Incumbency Advantages, and Campaign Finance, *The Journal of Politics* **70**: 681–699.

Meyer, M. A. (1991). Learning from Coarse Information: Biased Contests and Career Profiles, *Review of Economic Studies* **58**: 15–41.

Meyer, M. A. (1992). Biased Contests and MOral Hazard: Implications for Career Profiles, *Annales D'Economie et de Statistique* **25**: 165–187.

Milgrom, P. (1981). Good News and Bad News: Representation Theorems and Applications, *Bell Journal of Economics* **12**: 380–391.

Moloney, K. (2001). The rise and fall of spin: Changes of fashion in the presentation of UK politics, *Journal of Public Affairs* **1**: 124–135.

Morelli, M. and Van Weelden, R. (2011a). Ideology and Information in Policy Making. Forthcoming in Journal of Theoretical Politics.

Morelli, M. and Van Weelden, R. (2011b). Re-election Through Division. https://www.sites.google.com/richvanweelden/files/pandering.pdf.

Prior, M. (2006). The incumbent in the living room: The rise of television and the incumbency advantage in us house elections, *Journal of Politics* **68**(3): 657–673.

Robinson, M. (1981). The three faces of congressional media, *The New Congress*, American Enterprise Institute Washington, DC, pp. 55–96.

Rogoff, K. (1990). Equilibrium political budget cycles, *The American Economic Review* **80**(1): 21–36.

Rogoff, K. and Sibert, A. (1988). Elections and Macroeconomic Policy Cycles, *Review of Economic Studies* **55**: 1–16.

Smart, M. and Sturm, D. M. (2006). Term Limits and Electoral Accountability. CEPR Discussion Paper No. 4272.

Snyder, J. M. and Stromberg, D. (2010). Press Coverage and Political Accountability, *Journal of Political Economy* **118**: 355–408.

A Appendix

A.1 Proof of Proposition 1

We begin with two preliminary results. In Lemma 6 we show that if the cost function is convex, the message sent by the incumbent is nondecreasing in his type.

Lemma 6. *Given a reelection rule q, if $c(\cdot)$ is strictly convex, and $e_q(\cdot)$ is a best response to $v_q(\cdot)$, then the corresponding message $\tilde{\theta}_q(\cdot)$ is non decreasing in θ.*

Proof Let $\theta_1 < \theta_2$, and denote $\tilde{\theta}_q(\theta_i)$ by $\tilde{\theta}_i$ and $Pr_\epsilon(reelection|v_q(\cdot), \tilde{\theta}_i, q)$ by $P(\tilde{\theta}_i)$. Since $e_q(\cdot)$ (and therefore $\tilde{\theta}_q(\cdot)$) is a best response to $v_q(\cdot)$,

$$\begin{aligned}
\pi P(\tilde{\theta}_1) - c(\tilde{\theta}_1 - \theta_1) &\geq \pi P(\tilde{\theta}_2) - c(\tilde{\theta}_2 - \theta_1) \\
\pi P(\tilde{\theta}_2) - c(\tilde{\theta}_2 - \theta_2) &\geq \pi P(\tilde{\theta}_1) - c(\tilde{\theta}_1 - \theta_2)
\end{aligned}$$

Rearranging:

$$c(\tilde{\theta}_2 - \theta_1) - c(\tilde{\theta}_1 - \theta_1) \geq \pi(P(\tilde{\theta}_2) - P(\tilde{\theta}_1)) \geq c(\tilde{\theta}_2 - \theta_2) - c(\tilde{\theta}_1 - \theta_2)$$

Since the distance between the two sets of points is the same: $|(\tilde{\theta}_2 - \theta_1) - (\tilde{\theta}_1 - \theta_1)| = |(\tilde{\theta}_2 - \theta_2) - (\tilde{\theta}_1 - \theta_2)|$, the convexity of $c(\cdot)$ implies that $\tilde{\theta}_1 \leq \tilde{\theta}_2$. $\square$

In Lemma 7 we find sufficient conditions so that each voter's best response is a threshold rule.

Lemma 7. *If $\tilde{\theta}_q(\cdot)$ is increasing and $g(\cdot)$ satisfies the MLRP, then voter i's best response is a threshold rule:*

$$v_q(s, \eta_i) = \begin{cases} 0 & if \quad s < k_i \\ 1 & if \quad s \geq k_i \end{cases}$$

where k_i is determined by $E[\theta|s = k_i, e_q(\cdot)] + \eta_i = \theta_M$ whenever this equation has a solution, and $k_i = +\infty$ $(-\infty)$ if $E[\theta|s, e_q(\cdot)] + \eta_i < (>)$ θ_M for all $s \in \mathbb{R}$. Moreover, k_i is decreasing in the preference parameter η_i.

Proof If the message $\tilde{\theta}(\cdot)$ is strictly increasing in talent θ and the noise distribution $g(\cdot)$ satisfies the MLRP (meaning expected message is strictly increasing in the signal) then the conditional expectation of the talent must be increasing in the signal received by the voter [Milgrom (1981)], i.e., if $s_1 < s_2$ then $E[\theta|s_1, e_q(\cdot)] < E[\theta|s_2, e_q(\cdot)]$.

Moreover, since no information is revealed from the challenger, the expected talent of the challenger coincides with the mean of the talent distribution. Therefore, a voter with partisan position η_i supports the incumbent if and only if:

$$E[\theta|s, e_q(\cdot)] + \eta_i \geq \theta_M$$

Since the conditional expectation is increasing and continuous, if the equation $E[\theta|s = k_i, e_q(\cdot)] + \eta_i = \theta_M$ has a solution it has to be unique and voter i follows a threshold rule in which $v(s, \eta_i) = 1$ if and only if $s \geq k_i$. Finally, by the monotonicity of the expectation, k_i is decreasing in η_i. □

Now we prove Proposition 1. For any $q \in (\underline{q}, \overline{q})$, and any equilibrium $e^*_q(\cdot)$ and $v^*_q(\cdot)$, if $c(\cdot)$ is convex, Lemma 6 implies that $\tilde{\theta}^*_q(\cdot)$ is nondecreasing in theta. By the MLRP this implies that $E[\theta|s, e_q(\cdot)]$ is nondecreasing in s, and therefore $v^*_q(\cdot, \eta_q)$ is nondecreasing in s where $\eta_q = H^{-1}(1-q)$. If $v^*_q(\cdot, \eta_q)$ is constant, then the outcome of the reelections is independent of the signal and hence the incumbent exert zero effort, $e^*_q(\cdot) \equiv 0$. But then $\tilde{\theta}^*(\theta) = \theta$ is strictly increasing in θ and Lemma 7, together with the fact that $\underline{q} < q < \overline{q}$ imply that $v^*_q(\cdot, \eta_q)$ is not constant. Therefore $v^*_q(\cdot, \eta_q)$ must be a threshold rule with some threshold k_q. By the monotonicity of the expectation, $v(k_q, \eta_i) = 1$ for all $\eta_i \geq \eta_q$. Moreover, for all $s < k_q$, $v(s, \eta_j) = 0$ for all $\eta_j < \eta_q$. Therefore, the incumbent is re-elected if and only if $s \geq k_q$, which proves part (ii).

Finally, given a threshold k_q, the probability of reelection for an incumbent that sends message $\tilde{\theta}$ is $\Pr(\tilde{\theta} + \epsilon \geq k_q) = 1 - G(k_q - \tilde{\theta}) = G(\tilde{\theta} - k_q)$, where the last equality comes by the symmetry of the noise distribution. We can now write the expected payoff of the incumbent as:

$$V(\theta, e, q) = \pi G(\theta + e - k_q) - c(e)$$

then the (local) first and second order conditions for the optimal effort level, $e^*_q(\cdot)$ are:

$$\begin{aligned} \pi g(\theta + e^*_q(\theta) - k_q) &= c'(e^*_q(\theta)) \\ \pi g'(\theta + e^*_q(\theta) - k_q) - c''(e^*_q(\theta)) &< 0 \end{aligned}$$

Finally, totally differentiating the first order condition with respect to θ:

$$\frac{\partial e^*_q(\theta)}{\partial \theta} = \frac{-\pi g'(\theta + e^*_q(\theta) - k_q)}{\pi g'(\theta + e^*_q(\theta) - k_q) - c''(e^*_q(\theta))} > -1$$

where the last inequality follows by the convexity of $c(\cdot)$. This implies that the message sent by the incumbent, $\tilde{\theta}_q = \theta + e^*_q(\theta)$, is strictly increasing in his type. □

A.2 Formulas for calibration

Given the quadratic cost and the normal distributions, the equilibrium of the model is the following:

$$\begin{aligned} e_H = e_L = e^* &= \tfrac{1}{c\sigma_\epsilon}\phi(\tfrac{\delta}{\sigma_\epsilon}) \\ e_M &= \tfrac{1}{c\sigma_\epsilon}\phi(\tfrac{e_M - e^*}{\sigma_\epsilon}) \\ k^* &= \tfrac{1}{c\sigma_\epsilon}\phi(\tfrac{\delta}{\sigma_\epsilon}) \end{aligned} \tag{17}$$

where $\phi(\cdot)$ is the standard Normal density distribution.

The probability of winning for an incumbent is given by equation (10) and hence the difference in the probability of winning between the incumbent and the challenger is:

$$x = 2Pr(reelection) - 1 = (1 - 2p)\left(1 - 2\Phi\left(\frac{k^* - e_M}{\sigma_\epsilon}\right)\right) \tag{18}$$

where $\Phi(\cdot)$ is the standard Normal cumulative distribution.

Note that Lee (2008) computes the difference in the probability of winning between a marginal winner and a marginal loser. This avoids the problem of unobserved heterogeneity between winners and losers, if there is sufficient unpredictable noise in votes. Posterior differences between a bare winner and loser thus must be caused by the fact of winning or losing. In our model, all the politicians come from the same distribution of talents and therefore they are ex-ante identical and the difference in the probability of winning comes entirely from having been incumbent.

To compute the average vote share, note that given a signal s the share of voters that support the incumbent is $H(E[\theta|s])$. Hence the average vote share is given by:

$$AVS = \sum_{j\in\{L,M,H\}} Pr(\theta_j) \int H(E[\theta|s])\, g(s|\tilde{\theta}_j)ds \tag{19}$$

and the difference in the average vote share between the incumbent and the challenger is

$$y = AVS - (1 - AVS)$$