Agency Problems in Political Science

Thesis by
Saba Devdariani

In Partial Fulfillment of the Requirements for the
Degree of
Doctor of Philosophy

Caltech

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2022
Defended August 16, 2021

Saba Devdariani
ORCID: 0000-0001-5737-4052

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my committee members, Alexander V. Hirsch, Thomas R. Palfrey, Michael Gibilisco, and Luciano Pomatto for their invaluable advice continuous support during my PhD study. I am particularly indebted to my academic advisor, Alexander V. Hirsch, for his guidance and encouragement throughout my academic research and daily life.

My work in this thesis has significantly benefited from comments from Michael B. Gibilisco, Laural Doval, Gabriel Lopez-Moctezuma, Jean-Laurent Rosenthal, Lindsey A. Gailmard, Joanna Huey, Jacob Morrier,Yimeng Li, Francesco Lanfranchi, Alejandro Robinson-Cortes, Shiyu Zhang, as well as participants at the HSS Graduate Proseminar.

I also thank Kapáuhi Stibbard, Laura Flower Kim, and especially Laurel Auchampaugh, whose dedication as staff members made studying at Caltech a much easier and enjoyable experience.

Finally, I would like to express my gratitude to my parents, my brother, and my close friends. Without their tremendous understanding and encouragement the past few years, it would have been impossible for me to have completed my studies.

ABSTRACT

This dissertation consists of three chapters analyzing agency problems in political science. More specifically the role of different/additional information available to the principal or the agent.

In Chapter 2 we analyze the effect of the politician's knowledge of the external shock on his policy decisions on the unrelated issues. Elected politicians cannot control some external shocks, even if they can still anticipate their occurrence better than the general public. How can politicians use these types of anticipated external shocks to their benefit? How do they change their pandering incentives? And how does a rational voter incorporate these seemingly irrelevant external shocks in her voting decision? We build on the political accountability model of Canes-Wrone Herron, Shotts (2001), adding the ability to the voter to observe her utility, which is affected by an external shock. The shock is observed by the incumbent politician but not by the voter. Our analyses show that for high or low enough magnitude external shocks, a politician's ability to anticipate them eliminates his pandering incentives in equilibrium. For medium negative shocks, pandering could be a "gamble for resurrection," while for medium positive shocks, it acts as an "insurance" to guarantee the reelection. We show that both of these pandering regions emerge in equilibrium. The politician's knowledge of the shock, overall, decreases the voter's welfare in equilibrium.

In Chapter 3 we endogenize the information acquisition for the voter to study what types of policy decisions voters pay attention to, and why, and how the rational voter attention affect the behavior of politicians in office. We extend the Canes-Wrone, Herron, Shotts (2001) model of electoral agency to allow the voter to rationally choose when to "pay attention" to an incumbent's policy choice by expending costly effort to learn its consequences. When attention is moderately costly the voter generally pays more of it after the ex-ante unpopular policy than the ex-ante popular one. Rational attention may improve accountability by encouraging the politician to be truthful. In some cases it may also severely harm accountability both by inducing a strong incumbent to "play it safe" with a policy that avoids attention, or a weak incumbent to "gamble for resurrection" with a policy that draws it. Finally, rational attention can induce or worsen pandering but never "fake leadership".

Chapter 4 analyzes delegated information acquisition with a biased agent who also has private information about the state of the world. The information acquired is public and its informativeness increases with costly effort. Equilibria are characterized for two cases: when the agent decides an effort level and when a principal imposes formal requirements on it. The analysis demonstrates that even when the principal cannot commit to an arbitrary decision rule, he benefits from imposing formal requirements by getting as much public information as possible and correctly aligning the biased agent's incentives. In the optimal mechanism, the principal incentivizes the low type agent to truthfully reveal her private information by requiring a relatively low amount of costly effort, while the high private report has to be followed by the maximum effort in public signal.

TABLE OF CONTENTS

Acknowledgements iii
Abstract iv
Table of Contents vi
Chapter I: Introduction 1
Chapter II: External Shocks and Anticipatory Pandering 4
2.1 Introduction 4
2.2 The Model 8
2.3 Baseline Model: Incumbent does Not Observe ϵ 14
2.4 The Main Model: Anticipated External Shocks 18
2.5 Welfare Analysis 31
2.6 Conclusion 35
Chapter III: Voter Attention and Electoral Accountability 37
3.1 Introduction 37
3.2 Related Literature 41
3.3 The Model 45
3.4 Preliminary Analysis 46
3.5 Preliminary Results 57
3.6 Equilibrium Characterization 62
3.7 Voter Welfare 70
3.8 Conclusion 72
Chapter IV: Formal Requirements on Costly Information 76
4.1 Introduction 76
4.2 The Model 82
4.3 Preliminary Analysis: Policy Decision 84
4.4 Agent Decides Effort 86
4.5 No Commitment with Formal Requirements 92
4.6 Welfare Comparison 97
4.7 Conclusion 100
Appendix A: Chapter II 102
Appendix B: Chapter III 110
B.1 Preliminary Analysis 110
B.2 Equilibrium Characterization 115
B.3 Main Proofs 120
B.4 Voter Welfare 133
Appendix C: Chapter IV 139

Chapter 1

INTRODUCTION

This dissertation is organized into three chapters examining agency problems in political science. The first two chapters build on the existing political accountability model, exploring the consequences of additional information for the principal and the agent. Chapter 2 assumes the politician has information regarding upcoming external shock, while Chapter 3 endogenizes information acquisition by the agent about the policy success/failure. Chapter 4 investigates how the principal can use the formal requirement on costly public information to extract private information from the agent. The model is applied to the operation of federal advisory committees but can be generalized to more cases with a similar setup.

Some external shocks are beyond the politician's control, either because preventative measures are too costly or because they fall within the jurisdiction of other organizations. Even if these external shocks are unavoidable, politicians are frequently better at anticipating their effects than the general public. In Chapter 2, we analyze how knowing about the external shocks alters the politician's pandering incentive on unrelated policy issues, and how voters rationally incorporate these external shocks into their electoral behavior. We build on the political accountability model of Canes-Wrone, Herron, Shotts, 2001, adding the ability to the voter to observe their utility, which is affected by the external shock. We assume that the voter cannot disentangle the effect of the external shock from the effect of the policy choice.

We show that the incumbent's policy choice affects his reelection prospects through a direct channel—a different voting strategy for different policy choices—and an indirect channel—the policy choice affects the voting decision through the voter's utility. A combination of these two channels shapes strategic incentives for the incumbent politician. Our analyses show that, for most realizations of the external shock, a politician's ability to anticipate them eliminates his pandering incentives in equilibrium. The politician only panders after observing medium magnitude shocks, both positive and negative. However, the pandering incentives are different between these two cases. For medium negative shocks, pandering is a "gamble for resurrection," while for medium positive shocks, it acts as an "insurance" against a possible mistake. Welfare analysis shows that anticipation of external shocks eliminates some unnecessary pandering in equilibrium and might even improve the selection, but overall, decreases the voter's welfare in equilibrium.

Political accountability models frequently assume voters are ill-informed, providing them with little information other than observed policy choices. The only models that allow voters to learn about a policy's success or failure assume that this knowledge is revealed to them exogenously. Even if the media or government agencies "exogenously" generate such publicly accessible information, we argue that the volume of this information, as well as the limited amount of time and resources, forces the voter to prioritize and make a strategic decision on which information to pay attention to.

In order to address this issue, Chapter 3 examines what kind of policy decisions voters pay attention to and how this influences incumbent politicians' behavior. We extend the Canes-Wrone, Herron, Shotts, 2001 model of electoral agency allowing voters to rationally acquire costly information about the incumbent's policy success. Our findings indicate that the voter generally pays attention after the ex-ante unpopular policy. Rational attention may improve accountability by decreasing the incumbent's rewards for choosing an ex-ante popular policy, increasing her rewards for choosing an ex-ante unpopular one, or both. However, asymmetry in information acquisition can also harm accountability. Finally, rational attention can cause or exacerbate pandering (a bias toward the ex-ante popular policy), but never "fake leadership" (a bias toward the ex-ante unpopular policy). The latter phenomenon (as uncovered in Canes-Wrone, Herron, and Shotts, 2001) thus requires an asymmetry in voter learning across policies that derives from a process both separate from costly information acquisition, and sufficient to overcome its effects.

Chapter 4 analyzes delegated information acquisition with a biased agent who also has private information about the state of the world. The information acquired is public and its informativeness increases with a costly effort. Equilibria are characterized for two cases: when the agent decides an effort level and when a principal imposes formal requirements on it. The analysis demonstrates that when the agent decides the effort level, separation only exists for high enough cost. In separating equilibrium, the low type acquires no public information, while the high type acquires just enough public information to separate itself from the low type.

Even when the principal cannot commit to an arbitrary decision rule, he benefits from imposing formal requirements by getting as much public information as possible and correctly aligning the biased agent's incentives. In the optimal mechanism, the principal incentivizes the low-type agent to truthfully reveal her private information by requiring a relatively low amount of costly effort, while the high private report has to be followed by the maximum effort in public signal.

Chapter 2

EXTERNAL SHOCKS AND ANTICIPATORY PANDERING

2.1 Introduction

At the end of 2014, Georgia's national currency—the Georgian Lari (GEL) - started a sharp devaluation against the US Dollar. Apart from the major long-term negative macroeconomic consequences for the country, many Georgians also experienced an immediate hit to their financial well being. A significant part of the population had their income in domestic currency but bank loans in USD, making it challenging to repay their debt. The mentioned circumstances influenced voters' behavior: voters turned against the government, and the ruling party lost its supporters. In reality, there was very little the government could do to stop the devaluation of GEL for two reasons: (1) the USD started increasing its value relative to all currencies, not only the GEL, (2) the currency exchange rate and the inflation level was monitored by the National Bank of Georgia (NBG), which is an independent institution, functioning apart from political interference. So the voters held the government responsible for financial problems caused by external shocks that were not under the elected government's control. This is just one example of these widespread phenomena, and there is empirical evidence showing that voting behavior often depends on circumstances that affect the voters' well being but is not a consequence of the incumbent's political decisions. For instance, natural disasters often decrease incumbent's support,[1] and consequences of the federal government's policy are sometimes attributed to local politicians, etc.[2]

Even if the government could not do much to prevent currency devaluation, one could argue that they at least could see it coming. If executives could anticipate this negative external shock, it could affect their decisions about other relevant policy choices. An external shock essentially affects the incumbent's relative popularity, and if he can anticipate what the voter's predisposition is, he could certainly use it to advance his reelection goals. As for the voter, when she gets a noisy signal by observing her own utility, but cannot disentangle the effects of the external shock and an unrelated policy choice, she rationally incorporates it in her voting decision (even when the politician cannot control the effects of the external shock).

[1] Abney and Hill (1966), Achen and Bartels (2002, 2016).
[2] Sances (2017).

How anticipating different external shock levels would shape politicians' incentives, whether it would make them more truthful or more deceptive (pander) by hiding behind the external shock effects, is nontrivial. We construct a theoretical model to capture these strategic interactions and investigate equilibrium behavior.

We take a classical two-period pandering setup with a representative voter and add both external shocks to the voter's utility and the voter's ability to observe her own utility. Incumbent politicians can anticipate these external shocks (which may be positive or negative), which will have significant positive or negative consequences on the voter's well being. Even though such shocks might not be under the direct control of the local politicians (as in the currency devaluation example discussed above), the voters who observe their utility usually are not well informed enough to distinguish the external shock's effect and might misassociate it as a signal of the performance of the incumbent politician. Having prior information about the external shock thus enables politicians to anticipate the hostile or favorable mood of the electorate before making any policy decisions.

If the incumbent politician also possesses some expertise to make a policy choice, then he has a mechanism to manipulate voters' electoral decisions through pandering. Pandering is a well-studied phenomenon when the politician deviates from the optimal behavior with the sole purpose of winning over the electorate and retaining office. This paper analyzes the intersection of these two informational advantages by the incumbent politician: (1) an anticipated external shock that enables him to somewhat predict the voter's prior disposition, and (2) policy expertise that creates the possibility of winning over the electorate through pandering.

In this paper, we show that there are two main channels through which the incumbent's policy decision affects his reelection prospects in equilibrium:

- Direct effect: the voter gives different reelection criteria based on policy choice;
- Indirect effect: policy choice affects the utility of the voter and thus her voting behavior.

We show that when the external shock is high in magnitude (either positive or negative), neither of these two effects matter, and the incumbent's policy choice has absolutely no impact on his reelection probability. Therefore, the incumbent politician never panders when faced with significant external shocks. Conversely, if the shock is low enough in magnitude, the policy choice only affects reelection probability through its impact on the voter's utility. Since the private signal is informative and following it maximizes the policy's success, the incumbent never panders when the anticipated shock is low in magnitude.

There are only two possible scenarios where the anticipated external shock might encourage the incumbent to pander: positive and negative medium magnitude shocks. Pandering incentives are somewhat different between these two cases. When the incumbent politician is faced with a medium negative shock, he knows that he is most likely to lose office because of these unfortunate external circumstances (even if he chooses the correct but unpopular policy). The only possibility of retaining office is choosing the popular policy even if it goes against his private signal. Therefore, for these external shock values, pandering is a "gamble for resurrection" for an otherwise doomed incumbent. On the other hand, when a politician anticipates a medium positive shock, his reelection prospects are rather good. In fact, he might get reelected even after choosing the wrong but popular policy. Therefore, in this case, pandering is insurance in case of a wrong policy choice (following the wrong private signal). We show that pandering plays both of these roles in equilibrium. Moreover, in both scenarios discussed, pandering succeeds in its mission only when the incumbent's private signal is incorrect. This explains why the pandering equilibrium only arises when low-type politicians are sufficiently incompetent; otherwise, our model's only equilibrium is truthful.

We do the welfare comparison between the equilibria of the baseline and the main model, in order to show the effect that the politician's anticipation of the external shock has on the voter's welfare. Our analysis shows that the politician's ability to anticipate external shocks always weakly harms the voter. The difference in welfare is mainly due to the different equilibrium pandering levels as well as different voting rules for the second period incumbent. For certain realizations of the external shock, the politician's anticipation of external shocks benefits the voter by eliminating "unnecessary pandering" in equilibrium. For medium-low external shocks, this model generates even higher utility than the truthful equilibrium of the baseline model due to the better selection of the candidates for the second period. However, the politician's knowledge of the external shock harms the voter due to worse selection even for some realizations of the external shock where the equilibrium pandering level is decreased.

Multiple empirical papers and one recent theory paper by **?** investigate the link between electoral behavior and external shocks. This paper moves one step ahead by showing the effects of external shocks that are entirely out of control of elected politicians, whether it is because preventative measures are not possible, too costly, or under the jurisdiction of other institutions (like the exchange rate and the National Bank in our earlier example). We show that even in that case, the voter's behavior is rationally affected by the external shock since she cannot disentangle its effect on her own welfare from the effect of an unrelated policy choice. This is the first paper to concentrate on the politician's strategic incentives in such a setting, and its main contribution is to explain how an informed politician's incentives change with different anticipated external shocks and how these shocks affect their pandering behavior.

Literature Review

This paper is closely related to the voting literature and especially the works discussing the politician's pandering behavior (**????**). We borrow the definition and basic setup from **?**, who define pandering as a politician's behavior to go against the optimal action since "an executive's reelection prospects may depend on catering to voters who are potentially ill-informed about public policy." In the standard setup, the incumbent has better private information or "policy expertise" (**?**), and the voter's prior beliefs are biased towards one "popular" policy decision. These factors together create the possibility of pandering. Two main differences of our setup are that the voter can observe her own utility, and that this utility is affected by an external shock unrelated to the politician's action. This makes the voter's utility an imperfect signal of the incumbent's performance. Therefore, in our model, the voter has more information than just the policy choice: she can infer the politician's performance through her own observed utility. Moreover, besides policy expertise, the politician is also better aware of the external shocks, enabling him to use this information to his benefit.

Since the voter makes a voting decision based on her own utility (as well as the observed policy choice), this paper is also related to the literature on economic voting (**????**). Moreover, since the incumbent has no direct impact on the external shock but might be judged by its effect on the voter's utility, our paper also relates to multiple empirical papers examining the electoral effects of the external shocks and voter rationality. This literature usually concentrates on two types of shocks: natural disasters (**?????**) or international economic shocks (**???**). Our model is better suited to examining the second type of external shocks; for example, politicians can somewhat anticipate the upcoming global economic trends and their effect on the country's economic growth.[3]

While some papers show that exogenous shocks do not influence incumbents' re-election prospects (**?**), the main consensus in this literature is that a link between electoral behavior and negative external shocks does exist (**???**). As briefly discussed earlier, the closest work to our paper is a recent theory paper by **?** showing that external shocks "can affect electoral outcomes, even if voters are rational and have instrumental preferences." However, the main idea behind their results is that even though external shocks (like natural disasters) are not under the direct control of the incumbent, politicians can still "soften" the damage (with better emergency preparedness, infrastructure, etc.). Consequently, overall welfare gives the voter some indication of the incumbent's performance (and type). Like their model, the representative voter in this chapter rationally considers their welfare change since it contains information on the incumbent politician's performance, but the underlying mechanism is rather different. In our setup, politicians cannot influence the occurrence or the impact of the external shock. However, the representative voter's utility consists of both: the politician's policy performance and an unrelated external shock, therefore making it a noisy signal of incumbent's performance and their type. More importantly, their paper mainly concentrates on explaining the rationality behind the voter's behavior, while our paper studies how anticipated external shocks affect an incumbent politician's pandering incentives.

The remainder of this chapter is organized as follows: Section 2 introduces the formal model and shows some preliminary analysis; Section 3 discusses the baseline model without private information about the external shock and characterizes its equilibrium. Section 4 analyses the main model with external shock, introduces an equilibrium and discusses the players' main strategic incentives for different external shock realizations. Section 5 shows the welfare comparison between the baseline and the main models. Section 6 concludes.

2.2 The Model

Formal Description

We consider a two-period model with an election at the end of the first period. There are two candidates—an Incumbent (I) and Challenger (C)—and one representative voter (V). We assume that voters have common interests, and therefore they are

[3]It is harder to believe that politicians anticipate natural disasters and hide the information from citizens to use it for their electoral benefits.

reduced to one representative voter. In each period, there are two possible states of the world $\omega \in \{A, B\}$ and two possible policy choices $y \in \{A, B\}$. The voter's prior is that state A is more likely; specifically $P(\omega = A) = \alpha > 0.5$. This means that the voter who wants the policy to match the state is ex-ante inclined towards the policy choice A. Because the voter is ex-ante inclined towards A, we refer to A as the "popular" policy. In each period, the incumbent politician has to choose a policy $y \in \{A, B\}$. The voter always observes the policy choice.

Types

Politicians get private signals about the state of the world $s \in \{A, B\}$. There are two types of politicians based on the strength of the private signal they receive:

- A High-type politician (I=H) who knows the state with certainty, i,e., $P(s = \omega | I = H) = 1$.
- A Low-type (I=L) politician who gets a noisy but informative signal. If this type of politician wants to match the state, he will follow his signal. Formally $P(s = \omega | I = L) = q > \alpha$.

The voter does not observe the politician's type, and the prior belief that the politician (either incumbent or challenger) is a high type is $P(I = H) = p$.

Utilities

There is no discounting. As for the utilities, politicians are policy motivated but only if they are in office, i.e., in each period, only the incumbent politician gets utility:

$$u_p = \begin{cases} 1 & \text{if } \omega = y \text{ and in office} \\ 0 & \text{otherwise} \end{cases} .$$

The voter always wants to match the state. Her per-period utility is:

$$u_v = \mathbb{1}\{\omega = y\} + \epsilon,$$

where ϵ has a symmetric distribution with mean 0. One can interpret this as an external economic shock that is not affected by the incumbent's policy choice. For example, political instability or an economic recession in partner countries (with large amounts of trade or remittances) will usually negatively influence the country's economic performance.

The distribution of ϵ is common knowledge. The incumbent politician knows the exact realization of ϵ and therefore can use it to increase his chances of reelection. As already discussed, ϵ can be interpreted as an external shock on the country's economy of which the government has a clear understanding, but ordinary citizens merely feel its overall effect on their well being. We assume ϵ is drawn from the triangular distribution between $[-1, 1]$.[4] The probability density function (PDF) of ϵ is triangle :

$$f(\epsilon) = \begin{cases} 0 \ \ if \ \ \epsilon < -1 \ \ or \ \ \epsilon > 1 \\ -|\epsilon| + 1 \ \ otherwise \end{cases} .$$

The voter can see her own utility u_v, but does not know the realization of the external shock ϵ. Therefore, she can only form beliefs about the incumbent's performance, and consequently his type, based on her observed utility, her prior belief about the external shock ϵ, and the incumbent's policy choice.

Sequence of the Game

- Nature determines each politician's type, the first period state ω, the incumbent's signal s and the first period external shock ϵ;
- Candidates observe their own type, the signal about the state s and ϵ;
- The incumbent I chooses a policy for the first period y;
- The voter V observes the policy y and her own utility u_v;
- The voter V chooses to either reelect the incumbent I or elect the challenger C;
- Nature selects the state for period 2 ω' and the signal for elected candidate according to his type s';
- The second-state external shock ϵ' is realized;
- The officeholder observes signal s' and ϵ' and chooses the policy for the second period y';
- The voter V and the officeholder get their utilities.

[4] This assumption will be later relaxed to generalize the result for the distribution with full support.

Strategies

The incumbent's action in the first period and of the elected candidate in the second period is to choose a policy y for that respective period. Let $\sigma_s^I(\epsilon)$ for $s \in \{A, B\}$ and $I \in \{L, H\}$ denote probabilities of choosing the popular policy A in the first period for each type and signal of the politician and realization of the anticipated shock.

As for the voter, she has to choose which candidate to vote for in the election. This decision will also depend on the chosen policy and the realization of her utility. Let $V_y^{u_v}$ with $y \in \{A, B\}$ denote the probabilities of reelecting the incumbent after each possible information set of the voter (chosen policy and observed utility).[5]

Discussion

In this paper, we consider forms of external shocks that are entirely out of control of the incumbent politician. In some cases, the politician's preventative measures (which can also signal his competence) might soften the impact of external shocks. But in this analysis, we consider external shocks that are either too unpredictable or costly to prevent that even "good" politicians would not implement actions against them,[6] or prevention just does not fall in the jurisdiction of the incumbent politician (for instance if we model an election for local politicians and the shock is supposed to be dealt with on a federal level). This is done to eliminate any direct signaling about incumbent's type based on the magnitude of the shock.

For example, we think of external shock as some global economic trend (as in the case of USD revaluation in 2014, change in oil prices, etc.), military actions or political instability in a trading partner country, or even certain natural phenomena (like a heatwave affecting the agricultural sector). In that case, the government usually has better information than the general public and can somewhat predict an occurrence of external shocks due to the numerous resources such as advisers, corresponding agencies with field experts, etc. Even if it is too late for preventative measures or simply not under the incumbent politician's jurisdiction, the executive can respond to these predicted external shocks on the voter's utility while making other policy choices.

[5] We slightly abuse notation and use $V_{y=A}^{\epsilon}$ instead of $V_{y=A}^{u_v=\epsilon}$.

[6] In this paper, we do not model the prevention, but one could construct another model where the incumbent could "mute" the effect of external shocks by paying some cost ex-ante. If this cost is too high, and high magnitude shocks happen with low probability, the incumbent would never implement preventative measures and we would be back to our setup.

Another particular assumption of the model is regarding the type of politicians' utility. Politicians get utility of 1 if and only if they are in office and the chosen policy matches the state of the world. Following **?**, this type of utility is standard in the pandering literature and gives politicians two motives: (1) matching the state and (2) getting re-elected. Our goal is to show under what conditions each of these two factors is more important in guiding politicians' behavior. Another way to model politicians' utility would be making him policy-motivated and add fixed office-holding benefits. Note that both of those structures result in similar strategic incentives for the politician (with different weights). Choosing our model's utility structure makes calculations easier since politicians do not care about matching the state when they are out of the office. This is also a sufficient condition to eliminate unusual and unlikely equilibria where the low-type politician would rather lose the election to be replaced with the high-type executive in the second period.

We also assume a specific functional form for the PDF of the external shock ϵ. A robustness check reveals that for our results to hold, the PDF should be a single-peaked function.[7] Changing the expected value of the external shock shifts all the cutoffs in equilibrium accordingly. For instance, if the voter expects a positive external shock, her equilibrium utility cutoff points increase after both policy choices. A single-peaked PDF ensures that the observed utility carries consistent information regarding policy consequences. Since such functions make high magnitude external shocks less likely to occur, high negative utilities are attributed to policy failure with a higher probability (while high positive utilities are attributed to policy success). This results in the voter's cutoff reelection strategy in equilibrium. If the PDF does not satisfy the given property, the nature of equilibrium changes. For instance, if the external shock is uniformly distributed, the voter's observed utility carries no information about policy consequences. Consequently, the indirect effect of policy decision on reelection probability disappears and our results converge to **?**.

The incumbent and the challenger in our model are ex-ante identical as they have the same probability of being a high type. Adding an additional parameter to differentiate the incumbent and the challenger moves mixing regions in equilibrium, but essentially, the equilibria and strategic incentives do not change.

[7] The PDF should be strictly increasing below some value of the external shock and strictly decreasing above the same value.

Preliminary Analysis

In the second period, absent the reelection motives, the officeholder's sole desire is to match the state given his information set. In our model, the incumbent politician has two sources of information: his private signal s about the correct state of the world, and his private information about the realization of the external shock ϵ. Since even the signal of the low-type politician is informative ($q > \alpha$) and the external shock ϵ is uncorrelated with the true state of the world, an elected politician always follows their informative private signal in the second period. Given this politician strategy, the voter's second period expected utility (since $E(\epsilon) = 0$) for the high type is 1 and for the low type is q. Consequently, the voting behavior at the end of the first period will be entirely guided by the voter's posterior belief about the incumbent's type.

In the first period, the incumbent politician faces a tension between his desire to match the state in the first period and his reelection incentives. This choice is trivial for a high-type politician: he always follows his private signal in the first period since no reelection incentive is worth sacrificing the immediate first-period utility with certainty. Formally, in equilibrium $\sigma^H_{s=A} = 1, \sigma^H_{s=B} = 0$ for any ϵ.[8]

We further differentiate the regions of ϵ in equilibrium based on the behavior of the low-type politician. We borrow the definition of pandering and truthfulness from Canes-Wrone, Herron, and Shotts (2001).

Definition 1 *(i) We call a region of* ϵ *truthful if* $\sigma^H_{s=A} = \sigma^L_{s=A} = 1$ *and* $\sigma^H_{s=B} = \sigma^L_{s=B} = 0$ *for all values of* ϵ *in the region, (ii) we call the region of* ϵ *pandering if* $\sigma^H_{s=A} = \sigma^L_{s=A} = 1, \sigma^H_{s=B} = 0$ *and* $\sigma^L_{s=B} \in (0, 1]$ *for all values of* ϵ *in given region.*

In the truthful regions, both types of politicians follow their signal even when it contradicts the voter's prior belief. In a pandering region, a high-type politician follows his private signal, and a low-type politician follows his signal if it does not contradict his prior beliefs and he panders after seeing unpopular signal B. For simplicity, in later sections, we denote σ to be the probability of choosing the popular policy A by the low-type incumbent who receives unpopular signal $s = B$, i.e., $\sigma \equiv \sigma^L_{s=B}$.[9]

---

[8] Observe that the low-type politician always choosing the popular policy cannot be an equilibrium, since, in that case, the policy choice B would be a perfect signal for being the high type and ensure reelection, giving the low type strict incentive to follow her private signal when it indicates policy B.

[9] When the politician observes external shock, these strategies will be the functions of ϵ.

2.3 Baseline Model: Incumbent does Not Observe ϵ

First, we consider a baseline model where the incumbent does not know ϵ and has the same prior belief as the voter. Therefore, the incumbent makes a policy choice only based on his private signal s. The voter still observes her own utility and the policy choice.

The Voter's Problem

In the second period, all politicians follow their own signal, and therefore, the voter always reelects the incumbent if her posterior belief about the politician being a high-type is at least p. Moreover, we have already established that a high-type politician always follows his private signal in the first period.

The voter has two sources of information to update her posterior beliefs about the likelihood that the incumbent is high ability: her own utility u and the policy choice y. Using Bayes' rule, the voter's posterior beliefs μ_x^u for $x \in \{A, B\}$ are:

$$\mu_x^u \equiv P(I = H|u, y = x) = \frac{P(u|H, y = x)P(H|y = x)}{P(u|H, y = x)P(H|y = x) + P(u|L, y = x)P(L|y = x)}.$$

The voter reelects the incumbent if and only if this posterior belief is higher than p. We can think of this as a two-level updating of the voter's posterior belief. First, she updates based on the observed policy choice. If the politician is truthful, then choosing the popular policy A is more likely to come from the high-type politician, and therefore it increases her posterior belief ($P(I = H|y = A) > P(I = H|y = B)$). When the low-type politician starts to pander, the popular policy choice becomes a weaker indication of being a high type. For a certain pandering level, the two policy choices result in the same posterior beliefs (absent information from her utility).

The second part of updating is based on the observed utility of the voter. Observing the utility, the voter gets a signal about the probability of the state being matched, through which she updates her posterior belief about the incumbent being a high type. Fixing the policy choice y, when the voter observes her own utility u she knows one of these two scenarios occurred:

1) The incumbent politician matched the state. From the utility structure, this would imply $u = 1 + \epsilon \implies \epsilon = u - 1$;

2) The incumbent politician mismatched the state. Similarly, this would mean $u = 0 + \epsilon \implies \epsilon = u$.

This would mean that based solely on the observed utility, the voter's belief that the incumbent matched the state $P(\omega = y|u)$ equals $P(\epsilon = 1 - u)$. Given the prior belief about the external shock ϵ, for $u \in [0, 1]$,[10] the probability that the state was matched increases in u. Therefore, a higher utility is a stronger signal for a matched state, which in turn increases the voter's posterior belief about the incumbent ability. For $u = u_x^*$, the voter is indifferent between voting for either candidate. Thus, in equilibrium, after observing the policy choice $x \in \{A, B\}$, the voter reelects the incumbent when her utility is more than u_x^* and elects the challenger otherwise. A combination of all the incentives discussed above gives us the following lemma:

Lemma 2 *When the incumbent does not anticipate the external shock, the voter's strategy is a best response i.f.f. she reelects the incumbent when* $\left[y = A \text{ and } u > u_A^*\right]$ *or* $\left[y = B \text{ and } u > u_B^*\right]$ *and votes for the challenger otherwise.*

Figure 2.1 visually represents the form of the voter's strategy discussed in Lemma 2.

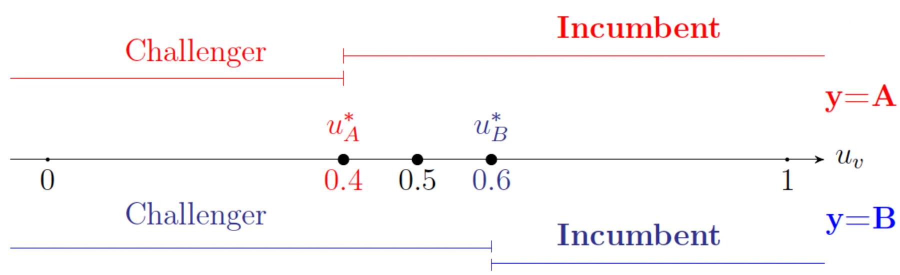


Figure 2.1: *This graph shows the voting strategy in truthful equilibrium when* $\alpha = 0.6$.

Note that, in a truthful equilibrium, these cut-offs only depend on α—the prior probability of the popular state.[11] In case of pandering we also have $u_A^* < u_B^*$—the voter is more "forgiving" when the incumbent politician makes the "popular" policy choice, which is the main driving force of pandering.

Corollary 3 *In equilibrium, the low-type incumbent always follows the signal that favors the popular policy* $s = A$.

[10] When $u > 1$, the voter knows with certainty that the state was matched since $\epsilon < 1$. Similarly, $u < 0$ is the perfect signal of mismatched state.

[11] Moreover, in that case $u_B^* = \alpha >= u_A^* = 1 - \alpha$.

The corollary follows directly from the discussion above. The incumbent politician always has two forces driving his behavior: (1) matching the state to get high first-period utility, and (2) maximizing his reelection prospects. When the low-type politician gets signal $s = A$, he strictly wants to follow it since both incentives work in the same direction: choosing policy $y = A$ guarantees a better reelection rule $u_A^* < u_B^*$ and a higher chance of matching the state.

The Incumbent's Problem and the Equilibrium

We have already established three facts about the incumbent's equilibrium strategy: (1) the second-period officeholder always follows his signal; (2) a high-type politician always follows his signal even in the first period; (3) a low-type politician always follows the popular signal $s = A$ in the first period. Now we discuss the strategic incentives of the low-type incumbent who receives the unpopular signal $s = B$. His expected utility from following this signal is:

$$EU_B^B = P(\omega = B|s = B)\left(1 + P\left(\epsilon > u_B^* - 1\right)P(\omega = s)\right) +$$
$$P(\omega = A|s = B)\left(0 + P\left(\epsilon > u_B^*\right)P(\omega = s)\right).$$

The first part of this expected utility represents the matched state: the incumbent gets first-period utility 1 as well as second-period expected utility $P(\omega = s)$ if he is reelected (with probability $P[\epsilon > u_B^* - 1]$). The second part is expected utility in case of a mismatched state: the incumbent gets utility 0 in the first period but still might get second-period utility if he is still reelected (with lower probability $P[\epsilon > u_B^*]$). Similarly, we can calculate the low-type politician's expected utility from choosing policy A after receiving private signal $s = B$:

$$EU_B^A = P(\omega = A|s = B)\left(1 + P\left((1 + \epsilon) > u_A^*\right)P(\omega = s)\right) +$$
$$P(\omega = B|s = B)\left(0 + P\left(\epsilon > u_A^*\right)P(\omega = s)\right).$$

One can clearly see where the pandering incentives come from the equations of expected utility. On the one hand, after seeing signal B the low-type incumbent wants to choose unpopular policy $y = B$ since it maximizes his chance of getting first-period utility. On the other hand, whether he matches the state or not, second-period reelection probabilities are higher if he panders (since $u_A^* < u_B^*$). The interaction of these two incentives guides the politician's equilibrium behavior.

We can calculate the equilibrium level of pandering $\tilde{\sigma}$ by setting $EU_B^B = EU_B^A$. Analysis shows that $\tilde{\sigma}$ is strictly decreasing in the competence level of the low-type incumbent q. This means that increasing the quality of the low-type politician decreases the probability of pandering. This effect is mainly due to the higher expected loss from disobeying the signal s. However, with higher q the voter's reelection cutoffs for different policies converge to each other. Moreover, since $EU_B^B - EU_B^A$ is increasing in σ, we can easily calculate values for q where a truthful equilibrium exists.

Proposition 4 *When the incumbent does not know the realization of the external shock ϵ, we have two possible equilibria:*[12]

1) *If $q > \tilde{q}$ only truth equilibrium exists;*

2) *If $q < \tilde{q}$, the low-type politician panders after observing $s = B$ with probability $\tilde{\sigma}$.*

In both types of equilibria, the voter's strategy is to reelect the incumbent i.f.f. ($y = A$ and $u > u_A^*$) or ($y = B$ and $u > u_B^*$).

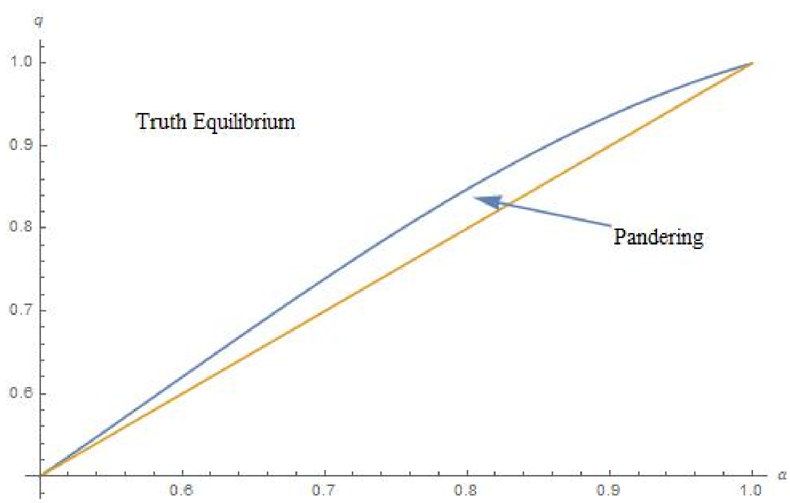


Figure 2.2: This graph shows regions of different types of equilibria when the voter observes her own utility. For a relatively high competent level of the incumbent, the cost of pandering is too high, and there exists truth equilibrium. For a relatively low level q, information of the voter and her voting strategy is not enough to make politicians behave truthfully, and low-type starts pandering after $s = B$.

[12]Proofs of the Propositions as well as the exact values of the parameters are given in the appendix.

In the pandering equilibrium, reelection cutoffs of the voter (u_A^* and u_B^*) depend on the prior probability of the popular state α as well as the competence level of the low-type politician q. The voter is still more likely to reelect the incumbent after a popular policy choice ($u_A^* < u_B^*$) creating pandering incentives. Moreover, the difference between the equilibrium cutoffs ($u_B^* - u_A^*$) increases with a more informative private signal of the low-type politician (higher q). This result is driven by the effect of q on the equilibrium pandering level. For more competent low-type politicians, the equilibrium pandering level decreases (eventually switching to the truthful equilibrium for $q > \tilde{q}$) since pandering becomes more costly. With a lower pandering level, the popular policy choice becomes a stronger signal for the politician's high type while the unpopular policy choice leads to lower posterior belief about the incumbent being a high type. Consequently, higher q decreases the equilibrium pandering level and causes divergence in equilibrium reelection cutoffs of the voter.

Next, we consider the main model where the incumbent anticipates the exogenous shock.

2.4 The Main Model: Anticipated External Shocks

For the main model's equilibrium, we consider a specific type of "cutoff" reelection strategy of the voter derived from the logic above.[13] This equilibrium strategy of the voter has the following form:

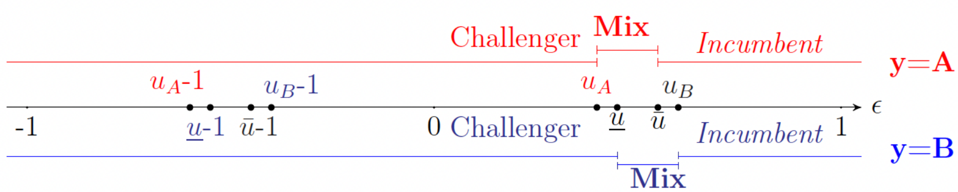


Figure 2.3: This graph shows the "cutoff" strategy of the voter in equilibrium, depending on the policy choice and her own utility.

If the voter sees the popular policy choice $y = A$ (red lines of the graph above), then:

- Elect the Challenger if the observed utility is low enough ($u \leq u_A$);
- Mix between voting for the Incumbent and the Challenger for an intermediate level of utility ($u_A < u \leq \bar{u}$);

[13] Note that we are describing equilibrium and we do not show the uniqueness yet.

- Reelect the Incumbent if the observed utility is high enough ($u > \bar{u}$).

Similarly, if the voter sees unpopular policy choice $y = B$ (red lines of the graph above), then:

- Elect the Challenger if the observed utility is low enough ($u \leq \underline{u}$);
- Mix between voting for the Incumbent and the Challenger for an intermediate level of utility ($\underline{u} < u \leq u^B$);
- Reelect the Incumbent if the observed utility is high enough ($u > u^B$).

Equilibrium

Similar to the baseline model, a truthful equilibrium with a simple cutoff reelection strategy exists only for sufficiently high competence levels of the low-type politician. Intuitively, when the low type's signal is more precise, his pandering cost is high enough to make him truthful. For lower levels of q, we get equilibrium where the voter uses a modified cutoff strategy with the mixing regions discussed above. Moreover, the incumbent politician panders if he faces medium positive or negative shocks, and he is truthful everywhere else. In the next section, we explain the rationale behind the equilibrium behavior of each player. Proposition 4 formally states the full equilibrium of this game with an anticipated external shock.

Proposition 5 *When the incumbent anticipates the external shock ϵ, there exists the following equilibrium:*

1) If $q > \bar{q}$ the politician is always truthful, and the voter reelects the incumbent i.f.f. ($y = A$ and $u > u_A = 1 - \alpha$) or ($y = B$ and $u > u_B = \alpha$);

2) If $q < \bar{q}$:

The Incumbent:

- *A low-type incumbent who receives private signal $s = B$ and anticipates a positive external shock panders with probability $\sigma_+^*(\epsilon) > 0$* [14] *when $\underline{u} < \epsilon < u_B$;*
- *A low-type incumbent who receives private signal $s = B$ and anticipates a negative external shock panders with probability $\sigma_-^*(\epsilon) > 0$* [15] *when $u_A - 1 < \epsilon < \bar{u} - 1$.*

[14] $\sigma_+^*(\epsilon)$ is increasing for $\underline{u} < \epsilon < \bar{u}$ and decreasing for $\bar{u} < \epsilon < u_B$.

[15] $\sigma_-^*(\epsilon)$ is increasing for $u_A - 1 < \epsilon < \underline{u} - 1$ and decreasing for $\underline{u} - 1 < \epsilon < \bar{u} - 1$.

Otherwise the incumbent is truthful.

The Voter: *After the popular policy choice* $y = A$*:*

- *The voter elects the challenger for low enough utility (*$V_A^* = 0$ *if* $u_v < u_A$*);*
- *The voter mixes between the incumbent and the challenger with probability* $V_A^* \in (0, 1)$ *for intermediate levels of utility (*$u_A < u_v < \bar{u}$*);*
- *The voter reelects the incumbent for high enough utility (*$V_A^* = 1$ *if* $u_v > \bar{u}$*).*

After the unpopular policy choice $y = B$*:*

- *The voter elects the challenger for low enough utility (*$V_B^{u*} = 0$ *if* $u_v < \underline{u}$*);*
- *The voter mixes between the incumbent and the challenger with probability* $V_B^* \in (0, 1)$ *for intermediate levels of utility (*$\underline{u} < u_v < u^B$*) ;*
- *The voter reelects the incumbent for high enough utility: (*$V_B^* = 1$ *if* $u_v > u_B$*).*

The exact equilibrium values (pandering and voting strategies) are given in the appendix. The graph below shows the equilibrium behavior of the low-type incumbent after observing unpopular signal $s = B$.[16]

[16]For illustrative purposes, the graph is constructed for the specific parameter values $\alpha = 0.6, q = 0.67$.

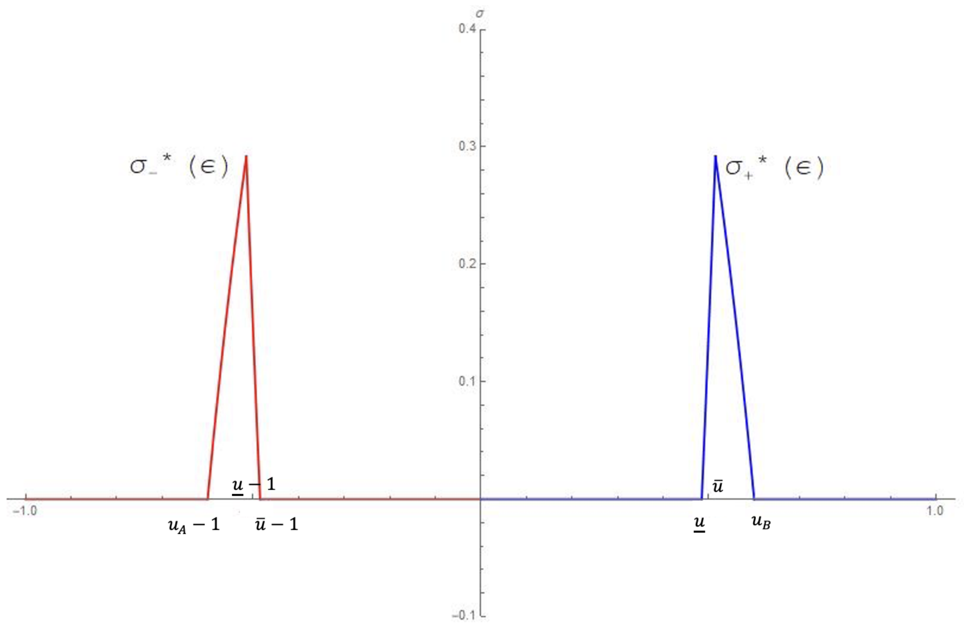


Figure 2.4: This graph shows the equilibrium level of pandering of the low-type incumbent after observing the unpopular signal $s = B$ depending on the realization of the external shock. The red line shows pandering after a negative shock (gamble for resurrection), and the blue line shows pandering after a positive shock (insurance). The graph is constructed for the specific parameter values $\alpha = 0.6, q = 0.67$.

Incumbent's Problem

For a high-type incumbent and a low-type with popular private signal $s = A$, nothing changes, and they are still truthful. Given the form of the voting strategy described above, we now discuss the problem of a low-type incumbent's who receives the unpopular private signal $s = B$, for different realizations of the external shock ϵ.

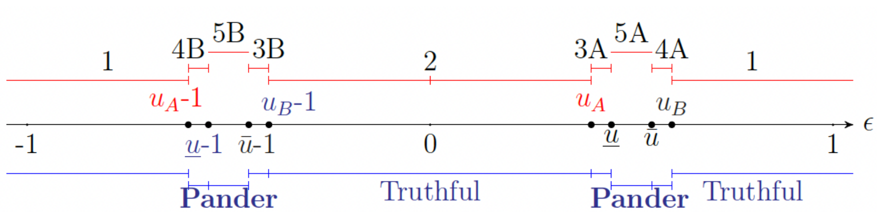


Figure 2.5: This graph shows equilibrium strategy of the low-type incumbent with unpopular signal $s = B$, depending on realization of the external shock ϵ.

If the low-type incumbent sees unpopular private signal $s = B$, his expected utility of following this signal and choosing policy B equals:

$$EU_B^B = P(\omega = B|s = B)(1 + V_B^{1+\epsilon}q) + P(\omega = A|s = B)(0 + V_B^{\epsilon}q).$$

If instead he panders by choosing the popular policy A, he gets:

$$EU_B^A = P(\omega = A|s = B)(1 + V_A^{1+\epsilon} q) + P(\omega = B|s = B)(0 + V_A^{\epsilon} q).$$

Therefore, the net expected benefit of following unpopular signal $s = B$ is:

$$\begin{aligned} EU_B^B - EU_B^A = &\underbrace{\Big(P(\omega = B|s = B) - P(\omega = A|s = B)\Big)}_{1} + \\ &\underbrace{\Big(P(\omega = B|s = B)V_B^{1+\epsilon} + P(\omega = A|s = B)V_B^{\epsilon}\Big)q}_{2} - \\ &\underbrace{\Big(P(\omega = A|s = B)V_A^{1+\epsilon} + P(\omega = B|s = B)V_A^{\epsilon}\Big)q}_{3}. \end{aligned}$$

The first part represents the first-period benefit of following signal $s = B$. The second part represents the future net benefits of following signal B, since q is the future expected utility of being reelected. The third part is the future benefit of pandering, and q is again the future expected benefit of being reelected. The difference between the second and the third part shows the second-period benefit of pandering for the low-type politician with the unpopular signal. This pandering benefit is especially large when the policy choice plays a rather important direct role for reelection. For example, the pandering incentives are high if the voter reelects the incumbent after the popular policy choice (but not after the unpopular policy choice) even with the mismatched state $V_A^{\epsilon} = 1$. Another such example is when the unpopular policy choice always leads to losing office even with matched state $V_B^{1+\epsilon} = 0$. These two cases drive the two distinct motives of pandering in the equilibrium we discuss below. Now, we consider the incumbent's problem in more detail for different regions of ϵ on Figure 2.5.

High Magnitude Shocks: Policy Chose Has No Impact on Reelection

When the incumbent politician anticipates high magnitude shock, either positive or negative (Region 1 in Figure 2.5), his reelection probability is predetermined: for large negative (positive) shocks, the incumbent always loses (retains) office whether or not the state is matched. For these values of ϵ, policy choice has no direct or indirect impact on the incumbent's reelection prospects. Since first-period utility is always higher when following the informative private signal ($P(\omega = B|s = B) > P(\omega = A|s = B)$), the incumbent is always truthful in this region.

Low Magnitude Shocks: Only Matching the State Matters for Reelection

When the politician foresees a low magnitude external shock (Region 2 in Figure 2.5), he only gets reelected if the state of the world is matched. These small shocks translate into utilities close to 0 and 1, giving a strong enough signal to the voter that the incumbent matched or mismatched the state, so that she votes solely based on this observed utility u_v. Therefore, policy choice has an only indirect effect on reelection probability through its impact on the voter's utility. Given the utility levels, different policy choices do not lead to different reelection behavior. Since even a low-type incumbent's private signal is informative, following it would therefore maximize both his first-period benefit and his second-period reelection probability. Therefore, this region is truthful in equilibrium.

Medium-Low Level Shocks Can Still Result in Truthful Behavior Even with Different Election Rules for Different Policies

For medium-low magnitude shocks (Regions $3A$ and $3B$ in Figure 2.5)[17] the policy choice has both direct and indirect effects on reelection probabilities. For positive shocks ($3A$), matching the state leads to reelection of the incumbent with certainty. However, if the state is mismatched, different policy choices lead to different electoral behavior: the voter reelects the incumbent after the popular policy choice $y = A$ with probability $V_A^\epsilon > 0$, while she never elects the challenger after observing the unpopular policy $y = B$; this creates pandering incentives for the low-type incumbent.

For positive medium-low magnitude external shocks ϵ, pandering is insurance against a possible mistake. When the low-type incumbent receives the private signal $s = B$ and follows it ($y = B$), he loses office if the state is mismatched. If he, instead, chooses the popular policy $y = A$, he may still be reelected if the state is mismatched (since $V_A^\epsilon > 0$). The reelection probability V_A' that would make the incumbent indifferent between following the unpopular signal or pandering in this region equals:

$$V_A' = \frac{(1+q)(P(\omega = B|s = B) - P(\omega = A|s = B))}{P(\omega = B|s = B)q}.$$

Since higher values of V_A discourage the incumbent from being truthful, being truthful will be a best response only for $V_A < V_A'$. We later compare this reelection

[17] Note that Regions $3A, 3B, 4A, 4B, 5A, 5B$ always exist since we have $1 - \alpha = u_A < \underline{u} < \bar{u} < u_B = \alpha$.

probability to the one from Region $4B$ to show that in equilibrium the incumbent is indeed truthful in Region $3A$.

For negative medium-low magnitude external shocks (Region $3B$ in Figure 2.5), the mismatched state leads to losing office after either policy choice. However, if the state is matched, the voter always reelects the incumbent after the popular policy choice $y = A$, and only sometimes reelects the incumbent after the unpopular policy choice $y = B$ (with probability $V_B^{1+\epsilon}$). This again creates different pandering incentives for the low-type incumbent compared to the previous case.

For these values of the external shock, pandering is a gamble for resurrection. In this region, the low-type incumbent who follows his unpopular signal almost always loses office (he is dismissed after the mismatched state and only sometimes reelected after the matched state with probability $V_B^{1+\epsilon} < 1$). If instead he panders and chooses the popular policy $y = A$, he guarantees himself reelection after the matched state. The reelection probability $V_B^{'}$ that makes the low-type incumbent indifferent between following the unpopular signal or pandering equals:

$$V_B^{'} = \frac{P(\omega = A|s = B)(1 + q) - P(\omega = B|s = B)}{P(\omega = B|s = B)q}.$$

Since V_B encourages the incumbent to be truthful, he will be truthful only for $V_B > V_B'$. Note that in this region $V_B^{1+\epsilon}$ is the same variable as V_B^{ϵ} in Region $4A$. We later compare $V_B^{'}$ to the reelection probability from Region $4A$ to show that in equilibrium the incumbent is indeed truthful in Region $3B$.

Medium-High Level Shocks Lead to Pandering in Equilibrium

Similar to the previous case, a medium-high positive shock (Region $4A$ in Figure 2.5) causes an incentive to pander as insurance against a possible mistake. Choosing the popular policy $y = A$ guarantees reelection, while choosing the unpopular policy $y = B$ might lead to losing office after the mismatched state. The reelection probability V_B^* that makes a low-type incumbent with the unpopular signal indifferent between choosing either policy is:

$$V_B^* = \frac{P(\omega = A|s = B)(1 + q) - P(\omega = B|s = B)}{P(\omega = A|s = B)q}.$$

$P(\omega = B|s = B) > P(\omega = A|s = B) \implies V_B^* > V_B^{'}$ means that any reelection probability that makes the incumbent pander after a negative medium-low shock from region $3B$ is not high enough to make him truthful after the corresponding positive medium-high shock from Region $4A$. Thus, Region $4A$ is a pandering region in equilibrium, while Region $3B$ is a truthful region.

For negative medium-high shocks (Region $4B$ in Figure 5), pandering is a gamble for resurrection. For these values of ϵ, choosing the policy $y = B$ leads to losing office with certainty. Choosing the popular policy $y = A$, might lead to reelection with probability $V_A > 0$ if the policy choice succeeds. The reelection probability V_A^* that makes the low-type incumbent indifferent between following the unpopular signal and pandering is:

$$V_A^* = \frac{P(\omega = B|s = B) - P(\omega = A|s = B)}{P(\omega = A|s = B)q}.$$

Note that $V_A' > V_A^* \implies (1 + q)P(\omega = A|s = B) - P(\omega = B|s = B) > 0$. This condition is always satisfied for $V_A^* < 1$. Therefore, if the incumbent panders after a positive medium-low external shock from region $3A$, the reelection probability V_A^* is higher than what makes him truthful after the corresponding medium-high negative shock from region $4B$. Since this reelection probability discourages the incumbent from being truthful, $3A$ is a truthful region and $4B$ is a pandering region in equilibrium.

Medium-Level Shocks Lead to Pandering with Mixed Reelection Probabilities after Both Policy Choices

When $\underline{u} < \epsilon < \bar{u}$ (Region $5A$ in Figure 2.5), the voter always reelects the incumbent if the state was matched regardless of the policy choice ($1 + \epsilon > u_B > \bar{u}$). If instead the state is mismatched, then the voter mixes by voting for the incumbent with probabilities V_A^ϵ and V_B^ϵ after the respective policy choices. When the low-type incumbent receives the unpopular signal $s = B$, if he follows it and chooses the policy $y = B$ he gets:

$$EU_B^B = P(\omega = B|s = B)(1 + q) + P(\omega = A|s = B)(0 + V_B^\epsilon q).$$

If instead he panders and chooses the popular policy $y = A$, he gets:

$$EU_B^A = P(\omega = A|s = B)(1 + q) + P(\omega = B|s = B)(0 + V_A^\epsilon q).$$

In equilibrium, since this is a pandering region, the low-type incumbent who observes the unpopular private signal $s = B$ is indifferent between choosing either policy, i.e., $EU_B^B = EU_B^A$. Simplifying this condition we get:

$$P(\omega = B|s = B)V_A^\epsilon - P(\omega = A|s = B)V_B^\epsilon = \frac{(1 + q)}{q}(P(\omega = B|s = B) - P(\omega = A|s = B)).$$

Observe that both reelection probabilities $(V_A^\epsilon, V_B^\epsilon)$ play a role in satisfying this indifference condition.

When $\underline{u} - 1 < \epsilon < \bar{u} - 1$ (Region $5B$ in Figure), the voter always dismisses the incumbent if the state was mismatched, regardless of the policy choice ($\epsilon < 0 < u_A > \underline{u}$). If instead, the state is matched, then the voter reelects the incumbent with probabilities $V_A^{1+\epsilon}$ and $V_B^{1+\epsilon}$ after the respective policy choices. When the low-type incumbent receives the unpopular signal $s = B$, if he follows it and chooses the policy $y = B$ he gets:

$$EU_B^B = P(\omega = B|s = B)(1 + V_B^{1+\epsilon}q) + P(\omega = A|s = B)0 = P(\omega = B|s = B)(1 + V_B^{1+\epsilon}q).$$

If instead he panders and chooses the popular policy $y = A$, he gets:

$$EU_B^A = P(\omega = A|s = B)(1 + V_A^{1+\epsilon}q) + P(\omega = B|s = B)0 = P(\omega = A|s = B)(1 + V_A^{1+\epsilon}q).$$

In equilibrium, since this is a pandering region, the low-type incumbent who receives the signal $s = B$ is indifferent between choosing either policy, i.e., $EU_B^B = EU_B^A$. Simplifying this condition we get:

$$P(\omega = B|s = B)V_B^{1+\epsilon} - P(\omega = A|s = B)V_A^{1+\epsilon} = \frac{1}{q}((P(\omega = A|s = B) - P(\omega = B|s = B)).$$

Note that $(V_B^{1+\epsilon}, V_A^{1+\epsilon})$ are the same reelection probabilities in this region as $(V_A^\epsilon, V_B^\epsilon)$ from Region $5A$. Indifference conditions of the low-type incumbent in Regions $5A$ and $5B$ define the values for reelection probabilities the voter uses in equilibrium. This is a rather intricate region with the incumbent pandering after both positive and negative external shocks. Such equilibrium structure is possible because of the overlapping regions of the voter's mixing strategies after the two policy choices (for $u \in (\underline{u}, \bar{u})$ the voter uses mixing reelection strategies after both popular and unpopular policy choices).

To sum up, for high levels of positive (negative) external shocks, the incumbent knows that he will always get reelected (dismissed), therefore, his policy decision has no direct or indirect effect on the voting behavior. This eliminates pandering incentives for the low-type politician for such external shock levels, making him truthful in equilibrium. The second truthful region of the equilibrium is after the low levels of external shocks. However, the rationale behind inducing the truthful behavior is very different from the previous case. In this region, the external shocks have a rather insignificant effect on the voter's utility, and matching the correct state becomes the decisive factor for reelection. Since following the incumbent's private signal maximizes the probability of matching the state, the low-type incumbent politician is truthful for these realizations of the external shock. Note that in this region policy choice has no direct effect on the voter's behavior—different policy choices lead to the same election decision for a fixed utility. Rather, the policy choice indirectly affects the voting strategy through its effect on the voter's utility.

Pandering incentives are only activated after medium-level external shocks (positive and negative). For medium positive external shocks, the incumbent anticipates a favorable predisposition of the voter and knows he is likely to be reelected. In this case, pandering almost guarantees the reelection and acts as an insurance. In contrast, for medium negative external shocks, the incumbent foresees an unfavorable position of the voter, thinking he is most likely to lose the office. Pandering in this region is a "gamble for resurrection." In our model's equilibrium, the pandering region emerges after both types of medium magnitude shocks.

The Voter's Problem

When the utility is large enough, $u_v > 1$, she knows with certainty that the correct policy was chosen (but she does not know the type of the incumbent). Since the incumbent and the challenger are drawn from the same pool and the matched state increases the voter's posterior belief that the incumbent is a high type, matching the state always leads to the reelection of the incumbent. Therefore, a voter who observes $u_v > 1$ always reelects the incumbent. If the utility is negative, $u_v < 0$, the voter knows that the state was mismatched. This is a perfect signal that the incumbent is a low type (only the low-type incumbent mismatches the state since the high-type politician knows the correct state with certainty and always follows his private signal). In equilibrium, there are two more regions of the voter's utility where the voter knows that the incumbent is always truthful (Figure 2.6):

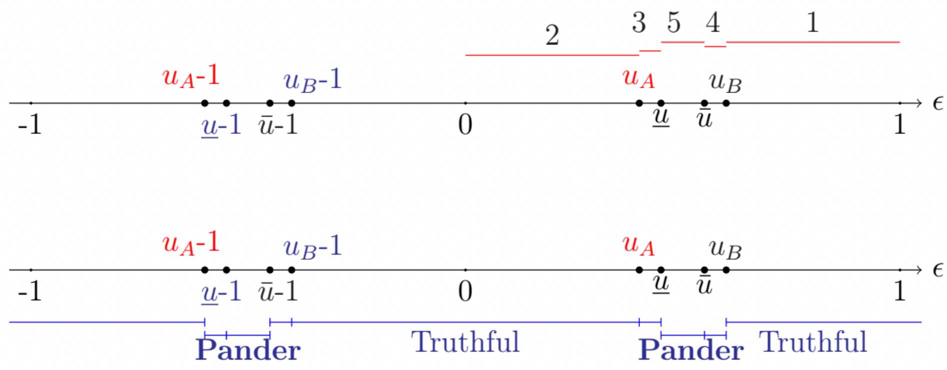


Figure 2.6: This graph shows regions for the equilibrium strategy of the voter, depending on the observed utility (above) and the equilibrium strategy of the low-type politician with unpopular signal $s = B$, depending on the realization of the external shock ϵ (below).

As we already discussed, the high-type politician is always truthful in equilibrium. The voter can infer the low-type incumbent's behavior from the observed utility and the policy choice.

For High or Low Positive Utility Levels, the Voter Always Knows the Politician Was Truthful

For high positive observed utility $u_B < u_v < 1$ (Region 1 in Figure 6), the voter knows that either (1) the state was matched and $\epsilon = u_v - 1 \in (u_B - 1, 0)$ or (2) the state was mismatched and $\epsilon = u_v \in (u_B, 1)$. In both cases, the low-type incumbent is truthful in equilibrium.

Similarly, for low positive observed utilities $0 < u_v < u_A$ (Region 2 in Figure 6), the voter knows that either the state was matched and $\epsilon = u_v - 1 \in (-1, u_A - 1)$ or the state was mismatched and $\epsilon = u_v \in (0, u_A)$. In either case, the incumbent is truthful in equilibrium.

With the knowledge that the low-type politician is truthful, the voter updates her posterior beliefs and votes for the incumbent when $P(H|u, y = B) \geq p$ i.f.f. $u > \alpha$. Intuitively, the voter sees two signals for the politician's type: the utility $u_v = u$ and the policy choice y. When the voter observes that $u > \alpha$ and knows that the incumbent was truthful, the utility is a stronger indicator of the high type of the politician than the policy choice. Consequently, even the unpopular policy choice leads to the reelection of the incumbent. Similarly, when the utility is low enough $u < 1 - \alpha$ and voter knows the incumbent was truthful, it is a strong indicator of the mismatched state and therefore the incumbent's low type. As a result, the voter elects the challenger even after the popular policy choice $y = A$.

For Medium-Low and Medium-High Positive Utilities, the Voter Does Not Know Whether or Not the Low-Type Incumbent Was Truthful

For medium-low positive utility $u_A < u_v < \underline{u}$ (Region 3 in Figure 6), the voter deduces that either the state was matched and the low-type incumbent was pandering,[18] or the state was mismatched and the incumbent was truthful.[19]

Based on this knowledge, the voter updates her posterior beliefs $P(H|u, y = A)$ and $P(H|u, y = B)$. For these observed utility levels, the voter is mixing between voting for the incumbent and the challenger after observing the popular policy choice $y = A$. The equilibrium level of pandering $\sigma^*_-(\epsilon = u_v - 1)$ in this region satisfies $P(H|u, y = A) = p$. Since in equilibrium the voter always votes for the challenger after observing the unpopular policy choice $y = B$ and the utility from this region, we also have $P(H|u, y = B) < p$. This condition derives the threshold $\underline{u}$ in equilibrium. Intuitively, when the observed utility increases, the equilibrium level of pandering $\sigma^*_-(\epsilon = u_v - 1)$ increases as well.[20] With higher levels of pandering, the probability that the low-type incumbent chooses the unpopular policy decreases, and therefore $P(H|u, y = B)$ increases. When the utility level surpasses $\underline{u}$, the voter no longer wants to vote for the challenger after observing $y = B$ ($P(H|u, y = B) > p$). Consequently, for higher utility levels we move to the region where the voter is mixing after observing either policy choice (Region 5).

Similarly, after observing medium-high positive utility $\bar{u} < u_v < u_B$ (Region 4 in Figure 6), the voter knows that either the state was matched and the low-type incumbent was truthful,[21] or the state was mismatched and the incumbent was pandering.[22]

[18] If the state is matched $\epsilon = u_v - 1$, in which case, $u_A - 1 < \epsilon < \underline{u} - 1$ and from the previous part, the incumbent is pandering.

[19] The state was mismatched and $\epsilon = u_v$, in which case, $u_A < \epsilon < \underline{u}$ and the low-type incumbent is truthful.

[20] The higher utility is a better signal for the high type of the incumbent and it should be balanced with the higher level of pandering to make the voter indifferent between voting for either candidate after the popular policy choice $y = A$.

[21] If the state was matched then $\epsilon = u_v - 1$, in which case, $\bar{u} - 1 < \epsilon < u_B - 1$ and from the equilibrium strategy, the incumbent is truthful.

[22] If the state was mismatched then $\epsilon = u_v$, in which case, $\bar{u} < \epsilon < u_B$ and the low-type incumbent is pandering with probability σ^*_+.

Based on this knowledge, the voter updates her posterior beliefs $P(H|u, y = A)$ and $P(H|u, y = B)$. In equilibrium, the voter is mixing between voting for the incumbent and the challenger after observing the unpopular policy choice $y = B$ and the utility level from this region. The equilibrium level of pandering $\sigma_+^*(\epsilon = u_v)$ is chosen to satisfy this indifference condition $P(H|u, y = B) = p$. Moreover, for the utility levels from this region, the voter elects the challenger after observing the popular policy choice $y = A$ in equilibrium. Therefore, we must have $P(H|u, y = A) > p$ for the equilibrium pandering level $\sigma_+^*(\epsilon = u_v)$. This condition is always satisfied for $u > \bar{u}$. Intuitively, with the lower levels of the utility in this region, the equilibrium pandering level increases.[23] When the pandering level increases, the popular policy choice $y = A$ becomes the weaker signal for the incumbent's high type, eventually decreasing $P(H|u, y = A)$ below p when the utility level reaches $\bar{u}$.

For Medium-Positive Utility, the Voter Knows in Equilibrium that the Low Type Politician Pandered

After observing medium-positive utility $\underline{u} < u_v < \bar{u}$ (Region 5 in Figure 6), the voter knows that either the state was matched and the low-type incumbent was pandering with probability $\sigma_-^*(\epsilon = u_v - 1)$[24] or the state was mismatched and the low-type incumbent was again pandering but with probability $\sigma_+^*(\epsilon = u_v)$.[25] Based on this knowledge, the voter's posterior beliefs about the incumbent being a high type after each policy choice are:

$$P(H|u, y = A) = \frac{f(u-1)p\alpha}{f(u-1)\alpha p + f(u-1)\alpha(1-p)(q + (1-q)\sigma_-^*) + (1-\alpha)f(u)(\sigma_+^* q + (1-q))(1-p)},$$

$$P(H|u, y = B) = \frac{f(u-1)p(1-\alpha)}{f(u-1)(1-\alpha)p + f(u-1)(1-\alpha)(1-p)q(1-\sigma_-^*) + \alpha f(u)(1-q)(1-\sigma_+^*)(1-p)}.$$

Since the voter uses mixing reelection strategy after both policy choices in this region, $\sigma_-^*(\epsilon = u_v - 1)$ and $\sigma_+^*(\epsilon = u_v)$ are chosen to make the voter indifferent between voting for the incumbent or the challenger after each policy choice. Therefore, these equilibrium pandering levels solve $P(H|u, y = A) = p = P(H|u, y = B)$.

[23] The lower utility level decreases the posterior belief of the voter about the incumbent's type, the higher equilibrium pandering level balances it out making sure $P(H|u, y = B) = p$.

[24] If the state was matched then $\epsilon = u_v - 1$, in which case, $\bar{u} - 1 < \epsilon < \underline{u} - 1$ and from the previous part the incumbent would pander with probability σ_-^*.

[25] If the state was mismatched then $\epsilon = u_v$, in which case, $\underline{u} < \epsilon < \bar{u}$ and the incumbent is again pandering but with probability σ_+^*.

2.5 Welfare Analysis

Next, we do the welfare comparison between the equilibria of the baseline and the main models in order to see how the politicians anticipation of the external shock affects the voter's welfare. In both cases, the voter is able to observe her own utility and the policy choice. However, in the main model, the incumbent politician can also anticipate the magnitude of the external shock. Even though the incumbent cannot change the impact of the external shock, he still uses it to advance his reelection prospects. From the structure of the equilibrium we discussed, it is not trivial to predict the overall impact of the incumbent anticipating the external shock on the voter's welfare.

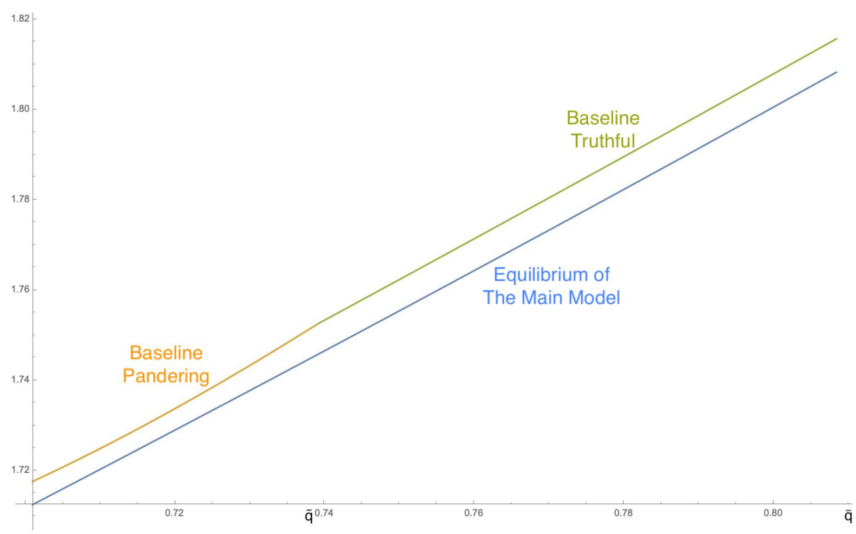


Figure 2.7: This graph shows the equilibrium level of the voter's welfare in the baseline and the main model, for different competence levels q of the low-type incumbent. The graph is constructed for the specific parameter values $p = 0.5$, $\alpha = 0.7$.

The graph above shows numerical calculations of the voter's equilibrium welfare for the two models for specific values of the parameters.

Conjecture 6 *The voter's equilibrium welfare weakly decreases with the incumbent's ability to anticipate the external shock.*

The relationship shown on the graph is general in the entire parameter space.[26] Overall, the voter is better off in the baseline model when the incumbent cannot

[26]The expressions for the expected welfare in the equilibria of two models are too complex to do the comparison by hand, but this result can be obtained using Mathematica.

anticipate the upcoming external shock. In order to better understand the forces driving this result, we discuss a few facts from the comparison of the two models.

Observation 1 *Everything else fixed, more pandering decreases the voter's welfare.*

There are two effects of pandering on the voter's welfare: the effect on first period utility, and the effect on selection for the second period. As pandering requires the incumbent to act against his informative signal, its effect on first-period utility is always negative. The higher the strength of the low-type incumbent's signal q, the greater is the expected damage that pandering incurs on the voter's first-period utility.

As for selection, the only advantage of pandering is when the failed attempt of pandering leads to the dismissal of the low-type politician from office and therefore to higher second-period expected utility for the voter. Observe that this only happens after the mismatched first period policy ("failed pandering"). Given the utility structure of the voter, she would never want to sacrifice her first-period utility for a potentially better politician in the second period. Therefore, if the voting strategy is fixed, more pandering is worse for the voter's welfare.

Given this observation, in order to better understand the welfare result stated above we next investigate how anticipating the external shock changes the pandering level, and what the impact of a more elaborate equilibrium voting rule in the main model is.

On one hand, when the incumbent anticipates the external shock, it eliminates "unnecessary"[27] pandering in equilibrium for certain realizations of the external shock ϵ. These are all the truthful regions in the equilibrium discussed in the previous chapter (either large or small shocks, both negative and positive sides).

Corollary 7 *For large enough or low enough external shocks in magnitude both positive and negative (Regions* 1 *and* 2 *in Figure* 8*), anticipating external shocks weakly increases the voter's expected welfare in equilibrium.*

[27] We call it unnecessary since, for these realizations of ϵ, pandering cannot increase the probability of reelection. Therefore, while the incumbent would still pander in the baseline model equilibrium (for certain values of q), he will always be truthful after anticipating such external shocks in the main model equilibrium.

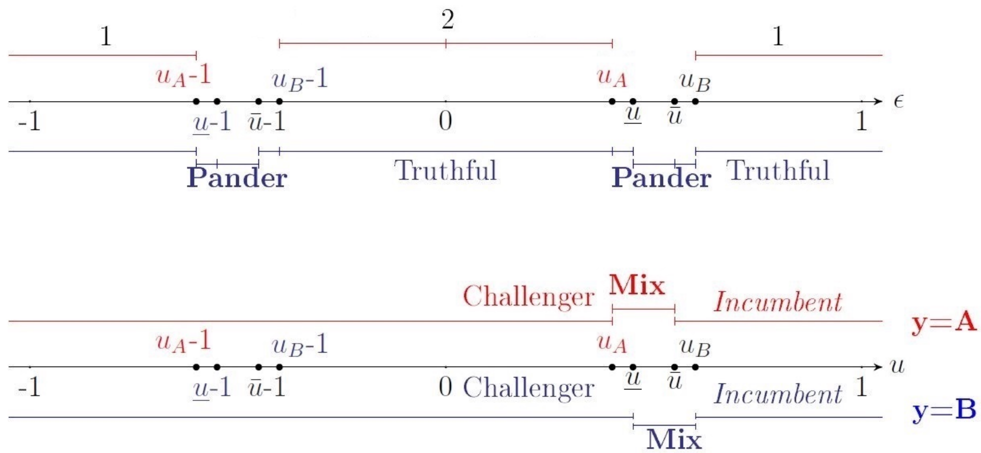


Figure 2.8: The voter benefits from the incumbent's knowledge of the external shock.

Trivially, for $q > \tilde{q}$ the baseline model has a truthful equilibrium so the welfare of the voter is the same in the two models. However, when the baseline model has a pandering equilibrium ($q < \tilde{q}$), anticipation of external shocks eliminates pandering for these realizations of ϵ. Moreover, for this region voting behavior in equilibrium is the same for the two models: always choose the challenger for high-negative shocks, always vote for the incumbent for large-positive shocks, and reelect the incumbent i.f.f. he matches the true state (based on observed utility of the voter) after low-magnitude shocks.[28] Therefore, it falls directly from the previous observation that the voter will benefit from the incumbent's knowledge of the external shock for these regions.

Earlier we mentioned that besides the pandering level, a different reelection rule also affects the expected welfare of the voter in equilibrium. For this reason, for medium levels of the external shock where the incumbent does not pander in equilibrium of the main model (Region 3 in Figure 9), the voter gets higher expected welfare even compared to the truthful equilibrium of the baseline model.

[28]The equilibrium cutoffs from the baseline model always satisfy $1 - \alpha = u_A < u_A^* < \underline{u}$ and $\bar{u} < u_B^* < u_B = \alpha$

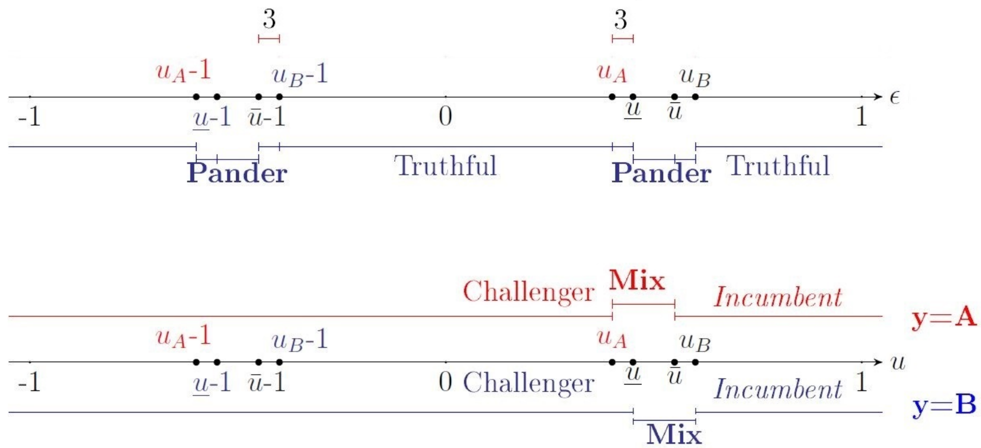


Figure 2.9: The voter benefits from the incumbent's knowledge of the external shock even compared to the truthful equilibrium of the baseline model.

Observation 2 *For medium realizations of the external shock that lead to a truthful behavior in the equilibrium of the main model (Region* 3 *in Figure* 9*), the voter's expected utility is higher even compared to the truthful equilibrium of the baseline model.*

This result is easy to see for positive-medium level shocks $\epsilon \in (u_A, \underline{u})$. When this type of shock is realized, everyone who matches the state is reelected in equilibria of both the main and the baseline models. When the state is mismatched (which only happens when the incumbent is a low type), in the truthful equilibrium of the baseline model, the low-type incumbent, who chooses the popular policy, is always reelected since $u_v = 0 + \epsilon > u_A$. However, for the same value of ϵ, in the main model equilibrium, the incumbent is only reelected with some probability $V_A^* < 1$. Therefore, for these levels of the external shock, the truthful low-type incumbent has less chance of reelection after choosing a popular policy in the main model equilibrium. This is the region showing the benefit of the anticipated external shock on the voter's welfare through reelection rule rather than a change in pandering level.

These are the only regions of the external shock that give an advantage to the anticipated external shocks. For other values of ϵ, the politician's knowledge of the upcoming shock decreases the equilibrium expected level of the voter's welfare. Even when the equilibrium level of pandering after a certain level of ϵ is lower in the main model equilibrium, worse selection leads to a lower level of expected welfare (for that region). This is easy to see for $\epsilon \in (\bar{u}, u_B^*)$.[29] For some of these values

[29] Where u_B^* is the equilibrium cutoff of the baseline model and $u_B^* < u_B = \alpha$.

of ϵ, the politician's knowledge of the external shock decreases the equilibrium pandering level but leads to a worse selection for the second period. The low-type incumbent who chooses the policy $y = B$ but fails to match the correct state loses office in the baseline model ($u_v = 0 + \epsilon < u_B^*$) but gets reelected with probability V_B^* in the equilibrium of the main model. This leads to the worse selection for the second period and overall it decreases the voter's welfare even with a lower level of pandering.

To sum up, the voter's equilibrium expected utility is higher for the main model for all truthful regions and lower for all pandering regions. The difference does not solely depend on the equilibrium pandering level but also on the reelection rule for the second period. By eliminating pandering and imposing a different election rule, the model with an anticipated external shock performs even better than the baseline model equilibrium for certain levels of external shock. In other regions, the politician's knowledge of the external shock harms the voter's welfare by increasing the pandering level or generating worse second-period selection even with lower pandering. Overall, despite some commonality in the voter's and the politician's motives (matching the state), when the incumbent can anticipate the external shock, the equilibrium expected welfare of the voter decreases.

2.6 Conclusion

In this chapter we show how anticipation of an external shock affects the politician's strategic decision to pander and the voter's welfare. Our findings suggest that, in equilibrium, the incumbent only panders when he expects a medium external shock. Comparing our equilibrium to the baseline model where the incumbent does not observe the external shock shows that giving information about ϵ to the politician eliminates pandering for some realizations of the external shock, but overall it harms the welfare of the voter.

There are two main channels through which the incumbent's policy choice affects his reelection prospects:

- Direct: the voter can make different voting decisions for different policy choices for a fixed observed utility;
- Indirect: the incumbent's policy choice affects the utility of the voter, which in turn is used in the voting decision.

When the shock is too big (either positive or negative), the incumbent's policy choice has no direct or indirect effect on the voting decision—the voter disregards the policy choice and both matched or unmatched states lead to the same voting behavior. A large enough positive shock guarantees the incumbent reelection whether or not he matches the state (similarly, a low enough negative shock leads to certain loss of office). Therefore, the incumbent's policy choice only affects his first-period utility and results in him being truthful.

For low-magnitude external shocks (either negative or positive), the incumbent's reelection outcome is not determined but rather depends on matching the state. Similarly, the voter's policy choice has no direct effect on reelection. Given the state is matched or mismatched, the voting behavior does not vary based on the observed policy choice. However, the indirect effect is still present in this case—he incumbent is reelected i.f.f. he matches the state. Even though the voter does not explicitly see the matched state, the observed utility is a good indicator of it for small shocks. Consequently, the incumbent's only objective (both for the first-period utility and the reelection prospect) is matching the state, which is more likely to be achieved following his informative private signal. As a result, the incumbent is always truthful when he sees a low-magnitude external shock.

Pandering incentives of the politician are activated only for moderate external shocks. In this region, both of the channels discussed above are present. Policy choice affects the utility of the politician by changing his reelection prospects, and it also has an immediate impact on the reelection probability even with fixed policy outcome (success or failure). In equilibrium, the incumbent panders after both moderate positive and negative shocks. In our equilibrium, pandering is insurance in case of a mistake under favorable external circumstances (positive ϵ) or a gamble for resurrection under unfavorable circumstances (negative ϵ).

Welfare analysis shows that the politician's ability to anticipate the external shock harms the voter. In the main model equilibrium, the voter benefits from eliminating the unnecessary pandering regions after either large or insignificant external shocks. However, the politician's anticipation of the external shock sometimes leads to higher levels of pandering for medium-level shocks, harming the expected utility of the voter. Even when the pandering level is decreased, the worse selection of the second period representative gives a disadvantage to the equilibrium of the main model. Overall, the voter's expected welfare in equilibrium weakly decreases with the politician's ability to anticipate the external shock.

Chapter 3

VOTER ATTENTION AND ELECTORAL ACCOUNTABILITY

Saba Devdariani
Alexander V. Hirsch

3.1 Introduction

The performance of the democratic process depends both on the information that voters possess and on how they use it. There is a long-standing debate in the elections literature about voters' competence to collect and process information (**?**)—whether they are well-informed, what kinds of information they possess, and how they use that information in their voting decisions. Following these debates, formal scholars have developed a variety of models to better understand how differing assumptions about voter information affect the accountability relationship. Scholars have examined models with voters who are entirely ignorant of incumbents' policy choices (**?**), who have information about their policy choices but not their consequences (**?**), and who can exogenously learn about policy consequences as well (**?**). They have also examined how the accountability relationship is influenced by information provided through strategic third-parties such as unbiased or biased media outlets (**????**) and democratic challengers (**?**). However, with the exception of **?**, existing models of political accountability assume —either implicitly or explicitly—that once information is made available to the voters, it is "free" for them to collect, interpret, and incorporate into their decisionmaking.

In reality, however, local and national news agencies cannot simply deposit information about policy performance directly into the minds of voters. Instead, their reporting must be actually be read (or viewed) and interpreted by voters to influence their decisionmaking. In addition to the news media, many academic centers as well as public and private research institutes produce reports that evaluate government performance in a large variety of policy areas on the local, state, and national levels. However, the target audience for such reports is typically an insular audience of policy professionals; for voters (or even reporters) to locate and digest this ostensibly-free and public information requires considerable time and effort. More generally, we argue that no matter how much free and even unbiased information is present in the public sphere, such information is never entirely free for voters to find and interpret. Instead, for such information to influence voters, they must (at least on some level) *choose* to spend some of their limited time and attention consuming and interpreting it. In this paper we seek to understand how this "attention constraint" affects voter behavior and democratic performance.

Our analysis is based on the canonical political-agency model of **?**. In this model, there is a representative voter who attempts to evaluate the incumbent's degree of *competence* at identifying effective policy alternatives. (There is no "partisan" element in the model, whereby voters and politicians may disagree intrinsically about the desirability of certain policy instruments or outcomes). In the baseline model, the voter evaluates the incumbent's competence by observing which policy he chooses. Even a relatively-incompetent incumbent is better informed than the voter about the likely efficacy of the available policies. However, one of the available policies is already perceived by the voters to be superior (that is, it is "popular"). Consequently, a less-competent incumbent fears that choosing the "unpopular" alternative—even if he privately believes it to be best—will be interpreted by the voters as a signal of his lack of competence and harm his electoral prospects. This leads him to sometimes pander and select the ex-ante popular policy even when he privately believes it to be mistaken.

To this model we add an ability for the voter to learn about the consequences of the incumbent's policy by paying *costly attention* after it is implemented. (An extension of the original model also examines the effect of information that may be *exogenously* revealed about the policy's consequences, but the main insights of that model are present when this possibility is absent, and our findings shed additional light on results when it is present). Thus, while our voter need not base his voting decision on the incumbent's policy choice alone, he must expend costly effort if he wishes to base his decision on something more. To simplify the analysis, we assume that a voter who pays costly attention perfectly learns the outcome of the incumbent's policy—that is, whether it succeeded or failed at achieving the intended goal. While strong, this assumption is intended to capture the conceptual opposite of existing analyses of voter information—that it is not information itself about incumbent performance that is scarce, but rather the attention required by voters to collect and process this information.

Forcing the voter to pay a cost to learn about the policy's consequences complicates the analysis—in our model, the voter must choose not only how to vote after each potential sequence of events, but also when to pay attention and learn more about the implemented policy. The first key insight of our analysis is that the disposition of the voter's attention is not neutral. Rather, a voter who is fully rational is looking *for* something specific when she chooses to pay attention. Specifically, she is looking for evidence that would reverse her current voting intention. Thus, if her current voting intention is to *retain* the incumbent, then she will only pay attention in order to find negative information about the incumbent's performance that would justify instead replacing him. Conversely, if her current voting intention is to *replace* the incumbent, then she will only pay attention to find positive information about the incumbent's performance that would justify instead retaining him. A key implication of this calculus is that the voter's willingness to pay attention will depend on the incumbent's policy in a particular way—it is not necessarily choosing the more or less popular policy that will garner the most attention, but rather the policy most likely to reveal an outcome that would reverse the voter's current voting intention.

Having established how and why rational voter attention will naturally be asymmetric across different policy alternatives, we next consider how this asymmetric attention affects the incumbent's incentive to disregard his private beliefs for electoral gain. In the original **?** model, the incumbent has an incentive to pander by choosing the ex-ante popular policy. This incentive derives from the fact that selecting an ex-ante popular policy signals expertise (under the twin presumptions that a high-ability incumbent is more likely to choose correctly, and the popular policy is more likely ex-ante to be correct). If the incumbent and the challenger have sufficiently similar initial reputations, pandering then becomes an effective strategy for securing reelection.

When the voter can choose whether to pay attention, however, this introduces a potentially-distinct rationale for the incumbent to disregard his private beliefs; to influence the level of attention paid by the voter. The reason is that an anticipated asymmetry in the voter's attention can lead a strong incumbent to want to "play it safe" by avoiding the policy that will draw more scrutiny, and a weak incumbent to want to "gamble for resurrection" by seeking out the policy that will draw it (e.g. **?**). In theory, an asymmetry in the voter's attention could bias the incumbent both toward or away from the initially-popular policy, depending on the incumbent's incentive to seek or avoid attention given her electoral standing. Closely related to this observation is a key finding in **?**; that when information about the *unpopular* policy is exogenously more likely to be revealed, a less-competent incumbent may sometimes pursue it even when he privately agrees that it is mistaken, gambling that his assessment is actually wrong and that the voters will then confuse his accidental success for competence. **?** term such a strategy “fake leadership.”

Our first main result is that when the voter’s information about policy consequences is filtered through rational attention, fake leadership cannot occur—even though rational attention is naturally asymmetric. The intuition is as follows. An incumbent who starts out strong has an incentive to avoid attention, while an incumbent who starts out weak has an incentive to seek it. At the same time, however, the voter’s willingness to pay attention after each policy depends on the possibility that this attention will reveal information that reverses her current voting intention. Thus, if the incumbent starts out strong enough that the voter is inclined to retain her, then the voter will be most willing to pay attention to the unpopular policy, because it is the one thought to be most likely to fail. Conversely, if the incumbent starts out weak enough that the incumbent is inclined to replace her, then the voter will be most willing to pay attention to the popular policy, because it is the one thought to be most likely to succeed. Thus, when the incumbent has an incentive to seek attention it is specifically pandering that will draw it, and when he has an incentive avoid attention it is *again* pandering that will deflect it.

We next consider which of the two policies will garner the most attention from a rational voter. Interestingly, we find that in general (though not always) it is the unpopular policy that will garner the most attention, despite the fact that it is the popular policy that is associated with pandering. Intuitively, the reason is that it is more difficult for the voter to “catch” a low-ability incumbent who is pandering than it is to “uncover” a high-ability incumbent who is exercising leadership. The former is so incompetent that he may accidentally achieve a policy success when he meant to pander, but the latter will always achieve a policy success when he meant to exercise leadership. While it is possible for the popular policy to garner more attention in equilibrium, this will only be the case under narrow conditions—specifically, when the incumbent is weak enough to be replaced after either policy absent attention, still strong enough to gain reelection if the voter uncovers a policy success, and the voter’s cost of attention is both low and falls in a small range. Under these conditions, the voter is relatively more willing to pay attention to the popular policy A because she thinks it more likely to reveal a success than the unpopular policy B, but she is not very willing to pay attention overall because a relatively strong challenger is already available.

We last consider how rational attention influences a low-ability incumbent’s pandering, and by implication the voter's welfare. We find that this effect depends both on the cost of attention and on the size of the competence gap between a high and low-ability incumbent. When the cost of attention is sufficiently low the voter will pay attention regardless of which policy the incumbent chooses. This will restore a low-ability incumbent's incentive to be truthful by tying is his electoral prospects to his policy successes. When the cost of attention is sufficiently high the voter will simply never pay attention, and so the incumbent's behavior will be unaffected its theoretical possibility.

However, if the cost of attention is *moderate*—so that the voter is inclined to pay attention after only one policy—then the competence gap between a high and low-ability incumbent determines whether the ability to learn about policy consequences ultimately improves or harms voter welfare. If the competence gap is not so large, then even asymmetric attention will be sufficient to reduce or eliminate a low-ability incumbent's incentive to pander; either because it "punishes" him with attention for choosing the popular policy, or because it "rewards" him with attention for choosing the unpopular one. However, if the competence gap sufficiently large, then asymmetric attention will introduce a new incentive to pander either to avoid attention (if the incumbent is strong) or to seek attention (if the incumbent is weak). This effect can both cause pandering that would not have occurred if the voter were unable to learn about policy consequences at all, and exacerbate pandering that would have already occurred. In fact, the harmful effect of asymmetric voter attention on accountability may be so severe that the voter would actually be better off were she unable to acquire *any* information about incumbent performance. A surprising implication is that voters may actually be harmed by the availability of public and unbiased information about incumbent performance once their "attention constraint" is taken into consideration.

3.2 Related Literature

Our model contributes directly to a now-large literature studying electoral accountability through the lens of principal-agent models. A substantial portion of this literature analyzes distortions in policymaking that are caused by forward-looking rational voters who lack the ability to commit ex-ante to their voting decisions (e.g. **?**, **?**, **?**, **?**). We build specifically on the canonical pandering model of **?**; in their work they discuss multiple real-life examples of politicians following popular opinion (i.e., pandering) because of re-election motives, and develop a theoretical model explicating the conditions under which pandering arises.

The main (indeed, only) difference between our model and **?** is that information about the success or failure of the incumbent's policy may only be revealed after an endogenous decision by the voter to acquire it. In modeling rational information acquisition by the voter, our work connects to several large literatures that collectively examine the effect of transparency and strategic information revelation in principal-agent relationships. These literatures are distinguished both by how such information is revealed—including generated exogenously, collected by the principal herself, or strategically revealed and/or collected by third party strategic actors—as well as the setting of the principal-agent relationship—including electoral representation, bureaucratic oversight, and the judicial hierarchy.

Within the study of electoral accountability in particular, earlier works sought to understand the effects of transparency by exogenously varying the process by which information was revealed to the voter—key works include **?**, **?**, and **?**. **?** argues that (exogenous) transparency about a politician's policy choice may decrease welfare by inducing pandering, but transparency about that politician's performance generally improves it. (**?** reaches a similar conclusion in a setting where incumbents differ in their preferences rather than their abilities.) In contrast with these works, **?** show that exogenous information about incumbent performance may also decrease voter welfare, if there is also an exogenous asymmetry in the *cost of choosing poorly*.

Subsequent works consider strategic acquisition and/or revelation of information by a variety of third parties, including biased and unbiased news agencies (**????**) and opposition parties (**?**). Similar to our findings, these papers suggest an ambiguous effect of information on accountability and voter welfare. **?** show that an unbiased media outlet that receives and reports on exogenous information about political performance can sometimes eliminate the incentive to pander. A fallible media, however, can also cause pandering that would not have occurred in its absence. **?** shows that media bias has an ambiguous effect on voters welfare, since an unbiased media improves selection but worsens accountability, while a biased media has the opposite effect. **?** shows that an opposition party can discipline the incumbent and increase accountability, but only if the opposition is sufficiently strong and policy-motivated. The main difference between these papers and ours is that we model endogenous information acquisition directly by the representative voter, instead of the third party media outlets with their own agenda. This enables us to examine the conditions under which the incumbent's policymaking is more or less likely to attract voter attention.

The only work of which we are aware that considers endogenous information acquisition by the voter in an electoral accountability setting is **?**. In that model, incumbents are differentiated by their preferences rather than their abilities. A key finding is that the voter pays too much attention to the incumbent's policy *choice* relative to the *consequences* of that choice, worsening accountability and voter welfare. However, these results largely derive from two modeling assumptions that do not hold in our setup: that the incumbent and the challenger are ex-ante identical, and that the information acquired by the voter is imperfect.

In examining information acquisition by the principal herself an agency relationship, our model also relates to several large literatures that span across political science, economics, public finance, and accounting studying "auditing" in principal-agent relationships. In these models, a principal (and sometimes other actors as well) can strategically acquire information about an agent's hidden actions or consequences thereof, and the acquisition of this information can influence the degree of moral hazard in the relationship. Such models have been applied most widely within political science to the study of bureaucratic politics (e.g. **?**, **?**, **?**, and **?**) and to the judicial hierarchy (see **?** for a review).

Notably, in many auditing models, an audit affects the agent's incentives predominantly by increasing the chance that the agent is "caught" deviating from the principal's wishes. For example, in the seminal judicial hierarchy model of **?**, a higher court (the "principal") only reviews cases decided by a lower-court (the "agent") when noncompliance is most likely, with potential reversal of the lower court acting as the punishment. In the setting of congressional oversight, ex-post audits are generally viewed as tools for detecting violations of legislative goals, whether it be through "police-patrols" (**?**,**?**, **?**) (that is, direct oversight by Congress), or "fire-alarms" (**?**) (that is, citizens and interest groups calling Congressional attention to deviant decisions). Similarly, in our model, better accountability can be induced with attention when it increases the risk that the agent will be caught pandering. However, accountability may also be improved through a different logic—it may make an agent more likely to be "found" having actually followed the principal's wishes despite having made a seemingly-bad policy choice.[1] In other words, rational attention may also improve accountability by functioning as a "reward" for

[1] Worth noting is that the absence of "auditing for compliance" in judicial hierarchy models also stems from the maintained assumption across this literature that "summary reversals" (in which a higher court can reverse a lower court decision without a costly rehearing of the case) are not possible.

choosing the unpopular policy.[2]

Our paper also speaks to a long literature examining voter competence. Early papers in this literature argued that voters do not possess enough information to effectively fulfill their electoral duties (**?**, **?**). Many authors argue that even when the voters are retrospective (**?**, **?**), or use other information such as party affiliation (**?**) and economic performance (**?**), they regularly misuse such information (**?**) or take into consideration irrelevant factors outside the control of policy-makers (**?????**. As pointed out by **?**, however, these debates have been "single-minded" in the sense that they do not typically take into consideration the *interaction* between the voters' information and the incentives of strategic politicians. Our paper is a further step in this direction—we investigate not only how voters rationally pay attention, but how that rational attention affects politicians' incentives to pander.

An additional literature to which our work relates is on rational inattention (RI), although we do not specifically adopt this technology to model the voter's information acquisition. The RI literature was started by Sims in a series of papers (**?**, **?**, **?**) on macroeconomic questions. Since then, similar tools have been used in finance (**?**, **?** **?**, etc.), labor economics (**?**, **?**, etc.), behavioral economics (**?**, **?**, **?**, **?**, etc.), and two-candidate elections models (**?**, **?**, etc.). A key finding of RI as applied to two-candidate election models is that voters will pay little attention when their personal stakes in an election are low. Moreover, in previous RI applications to elections the main factor determining a voter's endogenous information acquisition decision is her "pivot probability": if the voter believes that her vote will not be pivotal, she will not acquire information. In keeping with the political agency literature, however, our model features only a single representative voter so the question of pivotality does not arise. In addition, the voter already observes some information—the incumbent's policy choice—and her decision about acquiring information pertains to learning the consequences of that choice.

Finally, our paper also relates to **?**, who also study a voter's endogenous attention allocation but in an electoral competition model. In their model, the voter is distinguished both by her exogenous *interest* in politics and her endogenous *attention* to politics, and they show that the voter's attention only improves her welfare when she is moderately interested in politics. An additional difference between their work and ours is that the process of information revelation about candidates' choices in

[2] Similarly, see **?** for how rebates can be used to improve incentives to truthfully report income in a model of tax auditing.

their model is "two-sided"—it requires both costly attention from the voter as well as costly communication effort by the candidates.

3.3 The Model

We consider a two-period model with an election at the end of the first period. There are two candidates—an Incumbent (I) and Challenger (C) – and a representative voter (V). In order to avoid a pronoun confusion, we refer to the politicians as "he" and the voter as "she."[3] In each of two periods, nature draws a state of the world $\omega \in \{A, B\}$ that determines which of two potential policies $y \in \{A, B\}$ maximizes voter welfare.[4]

Information and Types The voter's prior belief $P(\omega = A)$ that the state is A in each period is denoted π. This is assumed to be strictly greater than $\frac{1}{2}$, implying that the voter is ex-ante inclined towards A; we therefore refer to A as the "popular" policy. Politicians, on the other hand, receive informative private signals about the state of the world $s \in \{A, B\}$. Specifically, each politician $j \in \{I, C\}$ may be either of high or low ability $\lambda_j \in \{H, L\}$. A high-ability politician ($\lambda_j = H$) learns the state with certainty ($P(s = \omega|\lambda_j = H) = 1$), while a low-ability politician ($\lambda_j = L$) receives a noisy but informative signal, where $P(s = \omega|\lambda_j = L) = q > \pi$. A politician's ability is his private information, and we denote the prior probability that the incumbent (challenger) is high ability as μ (γ)

Actions In each of two periods, the current officeholder chooses a policy $y \in \{A, B\}$, and this choice is observable to the voter. The "correct" policy in each period—i.e., the policy that maximizes voter welfare—is the policy that matches the state of the world ($y = \omega$). After the first period the voter chooses to reelect the incumbent or to elect the challenger. However, before making this decision (but after observing the politician's policy choice) the voter also chooses whether to the incumbent's "pay attention" to the policy decision ($\alpha \in \{0, 1\}$) by learning its consequences (i.e., her payoff), which costs c.

Utilities and Preferences. Players are assumed to have a common discount factor $\delta \in (0, 1)$. The voter only cares whether the officeholder in each period chooses the

[3] While the assumption of a representative voter is standard in the literature, it is more consequential in our model with costly information acquisition because the probability that an individual voter is pivotal in a large electorate is infinitesimal, but the cost of information acquisition is not.

[4] In a slight abuse of notation we do not superscript by period—throughout the analysis we make sure to clarify which period we are considering.

policy that maximizes her welfare. Specifically, in each period $U_V = \mathbf{1}_{\omega=y} - \alpha \cdot c$; i.e., the voter always wants the politician to match the state, and "paying attention" costs c.

Politicians are assumed to be policy motivated, but only if they are in office. That is, in each period a politician's utility is

$$U_j = \begin{cases} 1 & \text{if } \omega = y \text{ and } j \text{ is in office} \\ 0 & \text{otherwise} \end{cases}$$

This form of utility transparently combines two motives: (1) to maximize voter welfare, and (2) to get reelected.

Sequence of the Game The game proceeds as follows.

1. Nature determines each politician's type and reveals it her;
2. Nature determines the first period state of the world ω;
3. The incumbent I observes a first-period signal and chooses a first-period policy y;
5. The voter V observes the policy y, and chooses whether to pay attention $\alpha \in \{0, 1\}$:
   - If $\alpha = 1$ the voter V learns the her payoff U_V and pays cost c,
   - If $\alpha = 0$ the voter V learns nothing and pays no cost;
6. The voter V either reelects the incumbent I or elects the challenger C;
7. Nature selects the second-period state of the world;
8. The officeholder observes a second-period signal s' and chooses a policy y'.

The solution concept employed is Sequential Equilibrium.

3.4 Preliminary Analysis

In the last period, whoever holds office will follow his signal regardless of his ability (since $q > \pi$). Moreover, the voter will never choose to pay attention, since the only value of paying attention is to help decide whether to retain the current officeholder.

Incumbent's First Period Strategy In the first period, the incumbent politician I chooses a first-period policy y as a function of his private signal $s \in \{A, B\}$ and ability $\lambda_I \in \{L, H\}$. When doing so he may face a tension between his desire to match the state and his desire to get reelected. However, the only benefit of reelection in our model is the opportunity to maximize the voter's future welfare. Consequently, a high-ability politician will always strictly prefer to follow his first-period signal, since no increased likelihood of being able to maximize the voter's welfare "tomorrow" is worth sacrificing the voter's welfare for sure "today" (recall that $\delta < 1$). Correspondingly, we only introduce notation for the policy choices of a low-ability incumbent conditional on each possible signal; let σ_s denote the probability that a low-ability incumbent chooses policy A after signal $s \in \{A, B\}$.

Voter's First Period Retention After observing the incumbent's first period-policy, the voter forms an interim belief $\mu^x \in [0, 1]$ about the probability that the incumbent is high ability using Bayes' rule. This belief then determines the (interim) optimal probability of retaining the incumbent $v^x \in [0, 1]$ if she chooses not to pay attention. We term v^x the voter's *posture* toward the incumbent following policy x, since it reflects how favorably she treats an incumbent who chooses policy x if she chooses not to pay attention. If $v^x = 1$ (always reelect) we call the voter's posture *fully favorable*; if $v^x \in (0, 1)$ (sometimes reelect) we call it *somewhat favorable*; if $v^x = 0$ (always replace) we call it *adversarial*.

Voter's First Period Attention In our model, the voter must also choose whether or not to pay attention after observing policy x by paying c to learn her actual utility U_V (recall that $U_V = 1$ if and only if the incumbent's policy choice matched the state). However, since the incumbent's policy choice x is perfectly observable to the voter, learning that policy's consequences U_V is equivalent to learning the true value of the state ω. We therefore equivalently describe a voter who pays attention as one who learns the state, and let ρ^x denote the probability the voter pays c to learn the state after policy x.

In considering how a rational voter will choose to pay attention, observe that (as is standard in signaling models of electoral agency) the voter is unable to commit ex-ante to how she will respond to the incumbent's policy choice x, and therefore the probability she will pay attention ρ^x to each policy. This aspect of the model means that the voter may only rationally take into consideration how paying attention might improve *selection* (the probability that the second-period policyholder will be high ability) rather than *accountability* (how the incumbent uses her first-period information). In a broad sense, the voter's inability to commit to her attention decisions is what accounts for the potentially harmful equilibrium affects of (interim) rational attention.

The voter's inability to commit to attention also substantially simplifies the analysis. In particular, it implies that the voter will only pay costly attention when it might actually improve selection, which in turn is only the case if attention might reveal information that would persuade her to make a *different* retention decision than the one she intended (i.e., her posture v^x) based on policy alone. An immediate simplifying implication is that whenever the voter chooses to pay attention in equilibrium ($\rho^x > 0$), it must also be optimal for her to retain an incumbent who is revealed to have matched the state, and replace an incumbent who is revealed to have mismatched it (with at least one preference strict).[5]

The Incumbent's Problem

To analyze the calculus of a low-ability incumbent, observe that his utility from choosing policy $x \in \{A, B\}$ given whatever information $\mathcal{I}$ he has at the time of his decision is:

$$EU_{\mathcal{I}}^x = \underbrace{P(\omega = x|\mathcal{I})}_{\text{utility "today"}} + \underbrace{\delta q \left(\underbrace{(1-\rho^x)\, v_{\emptyset}^x}_{\text{no attention}} + \underbrace{\rho^x P(\omega = x|\mathcal{I})}_{\text{attention}} \right)}_{\text{utility "tomorrow"}}.$$

The contemporaneous benefit of choosing policy x is the possibility that it matches the state, which the incumbent believes will be the case with probability $P(\omega = x|\mathcal{I})$. The future benefit (discounted by δ) is the value of being reelected q (the probability that a low-ability incumbent's future signal will be correct) times the probability of reelection after choosing x. This probability, in turn, is equal to the voter's posture v^x if the voter chooses not to pay attention (with probability $1 - \rho^x$) and the probability $P(\omega = x|\mathcal{I})$ that x is correct if the voter does choose to pay attention (with probability ρ^x).

[5] More formally, let μ_ω^x and v_ω^x denote the voter's beliefs and retention probability after policy x and state ω. Since a high-ability incumbent always matches the state, an incumbent who mismatches the state must be low ability ($\mu_{\neg x}^x = 0$) and always replaced ($v_{\neg x}^x = 0$). It then follows that for attention to have value, the voter must strictly prefer to retain an incumbent who matches ($\mu_x^x > \lambda \rightarrow v_x^x = 1$); otherwise attention would not affect the voter's optimal retention decision.

Several features of $EU^x_{\mathcal{I}}$ are worth highlighting. First, the incumbent's rewards for choosing x are strictly increasing in his private belief $P(\omega = x|\mathcal{I})$ that x is correct. Consequently, a low-ability incumbent must be weakly more likely to choose a given policy x when his signal indicates that this policy is correct. This observation further implies that in equilibrium he may only sometimes disregard his private information in two ways: (i) by sometimes choosing the popular policy A even when his private information indicates that the unpopular policy B is correct ($\sigma_A = 1$ and $\sigma_B \in (0, 1)$) (which Canes-Wrone, Herron, Shotts (2001) term "pandering"), or (ii) by sometimes choosing the unpopular policy B even when his private information indicates that the popular policy A is correct ($\sigma_A \in (0, 1)$ and $\sigma_B = 0$) (which Canes-Wrone, Herron, Shotts (2001) term "fake leadership.")

Second, more voter attention after policy x makes the incumbent's utility from choosing x depend less on the voter's posture v^x, and more on that policy's actual quality. Thus, whether greater voter attention to policy x makes that policy more or less attractive to the incumbent overall depends on whether the incumbent's private belief $P(\omega = x|\mathcal{I})$ that x is correct is more or less favorable than the voter's initial posture toward x. A crucial implication is that *asymmetry* in the voter's attention can bias the incumbent's policy choice both toward or away from the policy indicated by his private signal.

The Voter's Retention Problem

In both the baseline version of the Canes-Wrone, Herron, Shotts (2001) model (henceforth CHS model) and in our model, the incumbent's incentive to follow her information is distorted by the voter's attempt to evaluate his ability via his policy decision.

To see this formally, observe that after seeing the incumbent's policy $y \in \{A, B\}$ the voter bases her retention decision on her posterior belief that the incumbent is high ability μ^y, and how it compares to her prior belief γ that the challenger is high ability. These posterior beliefs are calculated as follows. First, the probability a high-ability incumbent chooses policy A (or B) is simply the probability π (or $1-\pi$) that it is the correct policy. Second, the probability that a low-ability incumbent chooses policy A is $\sigma(\sigma_A, \sigma_B)$, where

$$\sigma(\sigma_A, \sigma_B) = (\pi q + (1-\pi)(1-q))\,\sigma_A + (\pi(1-q) + (1-\pi)\,q)\,\sigma_B.$$

Observe $\pi q + (1-\pi)(1-q)$ is the probability a low-ability incumbent receives a signal of A, and $\pi(1-q) + (1-\pi)\,q$ is the probability he receives a signal of B.

Then by Bayes' rule,

$$\mu^x(\sigma_A,\sigma_B)=\begin{cases}\frac{\mu\pi}{\mu\pi+(1-\mu)\sigma(\sigma_A,\sigma_B)} & \text{for } x=A\\ \frac{\mu(1-\pi)}{\mu(1-\pi)+(1-\mu)(1-\sigma(\sigma_A,\sigma_B))} & \text{for } x=B\end{cases}.$$

Using the above, it is simple to see that the voter's attempts to infer the incumbent's expertise from his policy choice is what gives the incumbent an incentive to pander. Specifically, when politicians are differentiated by their expertise and the voter deems policy A more likely to be correct ($\pi > \frac{1}{2}$), then she rationally believes that a high-ability incumbent's private signal is more likely to favor it. Consequently, if the incumbent is believed to always be "truthful"–that is, to always choose the policy indicated by his private signal—then a high-ability incumbent is *also* expected to more-frequently choose the popular policy A than a low-ability incumbent, i.e., $\pi > q\pi + (1-q)(1-\pi) = \sigma(1,0)$. Thus, choosing A (B) will be interpreted as a favorable (unfavorable) signal about the incumbent's ability, i.e., $\mu^A(1,0) > \mu > \mu^B(1,0)$.[6]

Pandering Absent Attention

To clarify how this effect can generate pandering in equilibrium and also establish a baseline against which to compare the model with rational attention, we conclude this section by revisiting the equilibrium of the CHS model with no attention ($\rho = 0$).

Let $\bar{\mu}^x$ denote the voter's posterior belief $\mu^x(1,0)$ about the incumbent's ability following policy x when she believes a low-ability incumbent to be truthful. Now first suppose that the voter's inferences when she believes the incumbent to be truthful are not strong enough to influence her retention decision, either because the incumbent is beginning very weak relative to the challenger ($\gamma \le \bar{\mu}^B$) or very strong ($\gamma \ge \bar{\mu}^A$). In these cases, a low-ability incumbent will clearly have no electoral incentive to pander, and the unique equilibrium will be truthful. Next suppose that the voter's inferences are sufficiently strong to influence her retention decision, i.e., $\gamma \in (\bar{\mu}^B, \bar{\mu}^A)$. Then a low-ability incumbent who privately observes signal $s = B$ faces a tradeoff between following his signal and getting reelected. Applying our previous characterization of a low-ability incumbent's expected utility $EU_{\mathcal{I}}^x$ from policy choice x given information $\mathcal{I}$ yields an incentive to pander after signal B if

[6] This "expertise pandering" effect is distinct from the "preference pandering" effect exhibited in models like **?**, in which the voter is uncertain not about the incumbent's expertise, but about whether his preferences match her own.

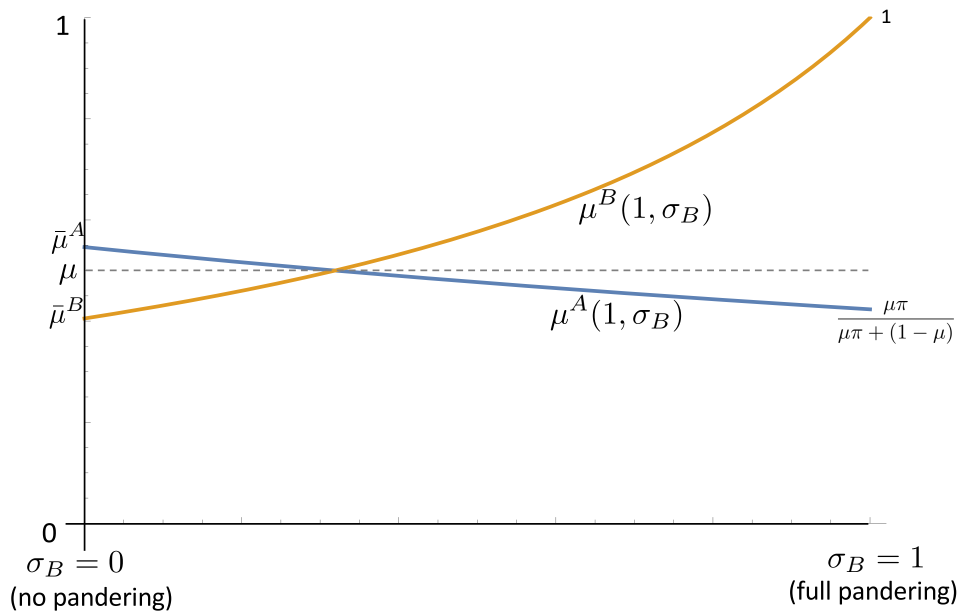


Figure 3.1: Posteriors after each policy as a function σ_B.

only if:

$$\delta q > P(\omega = B|s = B) - P(\omega = A|s = B),$$

or if the net future benefit δq of reelection exceeds the net current benefit of following the signal. It is straightforward to show that this inequality will hold (and thus that equilibrium must involve pandering) if and only if the quality q of a low-ability incumbent's information is below a unique threshold $\hat{q} \in (\pi, 1)$ that solves the equality:

$$\delta\hat{q} \cdot ((1-\pi)\hat{q} + \pi(1-\hat{q})) = \hat{q} - \pi. \tag{3.1}$$

To complete the characterization, let $\hat{\sigma}_B^x(\gamma)$ denote the unique probability of pandering that makes the voter indifferent between retaining and replacing the incumbent after observing policy x (i.e., $\mu^x\left(1, \hat{\sigma}_B^x(\gamma)\right) = \gamma$). (Recall that the baseline CHS model cannot exhibit fake leadership, so in equilibrium the incumbent will always follow a signal of A, i.e., $\sigma_A = 1$). As depicted in Figure 3.1, it is easily verified that (i) $\mu^x(1, \sigma_B)$ is strictly increasing (decreasing) in σ_B when $x = A$ (B), (ii) $\hat{\sigma}_B^x$ is well-defined when $\gamma \in \left[\bar{\mu}^B, \bar{\mu}^A\right]$, and (iii) there exists a unique $\hat{\sigma}_B(\gamma)$ satisfying $\sigma(1, \hat{\sigma}_B(\gamma)) = \pi$ (where a low and high-ability incumbent are equally likely to choose A), and at this pandering level $\mu^A(1, \hat{\sigma}_B(\gamma)) = \mu^B(1, \hat{\sigma}_B(\gamma)) = \mu$ (policy choice is uninformative to the voter). Equilibrium is then as follows.

Proposition 8 *Let σ_N^* denote the equilibrium level of pandering in the CHS model absent voter attention ($\rho = 0$). If a low-ability incumbent begins far ahead of or behind the challenger ($\gamma \notin \left(\bar{\mu}^B, \bar{\mu}^A\right)$),* ***or*** *if his information is sufficiently high quality ($q > \hat{q}$), then he is always truthful. Otherwise he must sometimes pander.*

- *If he is ahead of the challenger ($\gamma \in \left(\bar{\mu}^B, \mu\right)$) he panders with probability $\sigma_N^* = \hat{\sigma}_B^B(\gamma)$. The voter always reelects after A ($v^A = 1$) but only sometimes after B ($v^B \in (0, 1)$).*
- *If he is behind the challenger ($\gamma \in \left(\mu, \bar{\mu}^A\right)$), then he panders with probability $\sigma_N^* = \hat{\sigma}_B^A(\gamma)$. The voter sometimes reelects after A ($v^B \in (0, 1)$) but never after B ($v^B = 0$)*
- *If he is even with the challenger ($\gamma = \mu$), then he panders with probability $\sigma_N^* = \hat{\sigma}_B(\gamma)$, and there are a continuum of equilibrium voter retention probabilities $(v^A, v^B) \in [0, 1]^2$.*

Figure 3.2 depicts the type of equilibrium that prevails in the CHS model as a function of the challenger's reputation γ (on the x-axis) and the quality of a low-ability incumbent's information q (on the y-axis). A low-ability incumbent will be truthful either if policy choice is an insufficiently strong signal to influence the voter's retention decision, or if the quality of the incumbent's information makes pandering too costly. Otherwise, equilibrium must involve some pandering. In a pandering equilibrium the voter must "mix" by sometimes reelecting the incumbent after one of the two policies, but which policy this is depends on the incumbent's initial strength. If the incumbent is initially stronger than the challenger ($\gamma < \mu$), then the equilibrium level of pandering $\hat{\sigma}_B^B(\gamma)$ induces the voter to always retain the incumbent after policy A and sometimes retain him after policy B. If the incumbent is initially weaker than the challenger ($\gamma > \mu$), then the equilibrium level of pandering $\hat{\sigma}_B^A(\gamma)$ induces the voter to always replace the incumbent after policy B and sometimes replace him after policy A. Finally, equilibrium pandering is maximized (as a function of γ) when the incumbent and challenger begin equally matched ($\mu = \gamma$); pandering decreases and eventually vanishes as the challenger either becomes stronger or weaker than the challenger.

The Voter's Attention Problem

The distinctive feature of our model relative to CHS is that the voter need not rely only on what she can infer from the incumbent's policy choice; she may also expend

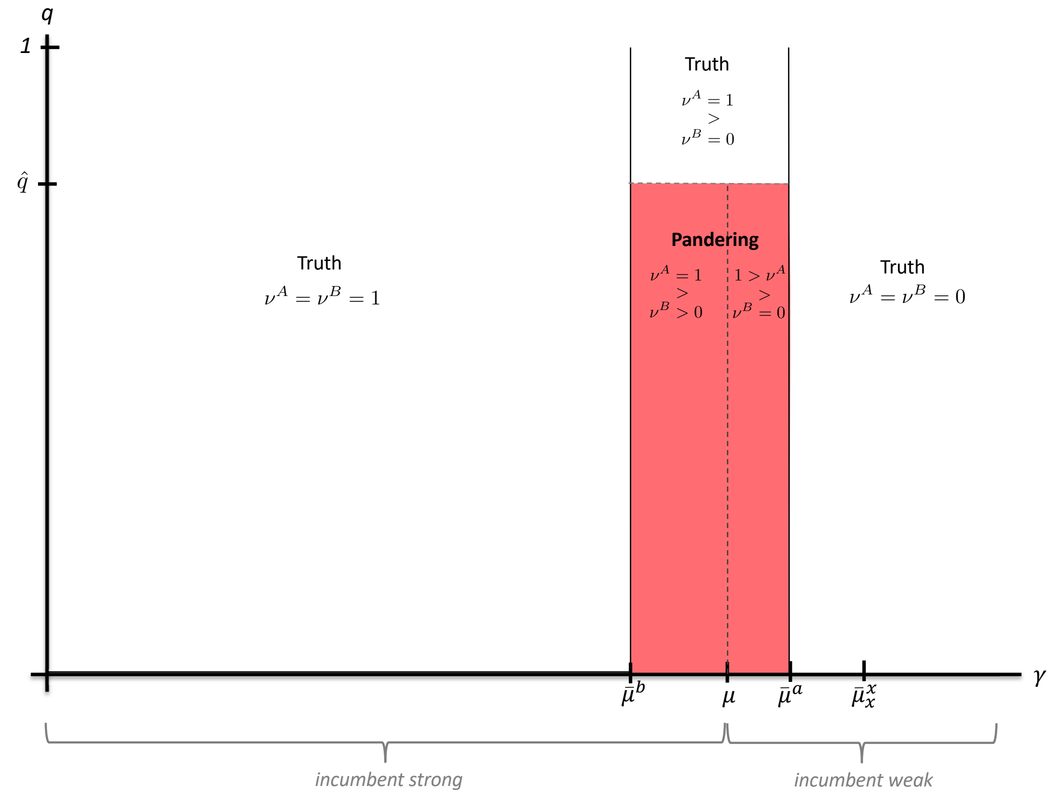


Figure 3.2: Canes-Wrone, Herron, Shotts (2001) equilibrium.

costly attention to learn the actual consequences of that policy (i.e., the state ω) in order to better evaluate the incumbent's ability. How the voter rationally allocates her attention after each policy choice, and how this rational allocation affects the incumbent's behavior, is the focus of our analysis. (Throughout this section we will temporarily suppress notation that explicitly indicates the dependence of the voter's beliefs and best responses on a low-ability incumbent's strategy (σ_A, σ_B)).

To begin the analysis, first let μ^x_ω denote the voter's posterior about the incumbent's ability after the incumbent chose policy x and attention reveals the state to be ω. If attention reveals that the incumbent did not match the state, then the voter infers that he is definitely low ability ($\mu^x_{\neg x} = 0$), since a high-ability incumbent both receives a perfect signal and always follows it.[7] Alternatively, if attention reveals that the incumbent did match the state, then the voter infers that he is high ability with

[7] Note that if a low quality incumbent always chooses A, then policy B being revealed to mismatch is off-equilibrium path, and the stated beliefs require the application of sequential equilibrium.

probability

$$\mu_x^x = \Pr(\lambda_I = H | y = x, \omega = x) = \frac{\Pr(y = x|\omega = x, \lambda_I = H)\Pr(\lambda_I = H)}{\Pr(y = x|\omega = x)}$$
$$= \frac{\mu}{\mu + \Pr(y = x|\omega = x, \lambda_I = L)(1-\mu)},$$

where $\Pr(y = x|\omega = x, \lambda_I = L)$ denotes the probability that a low-ability incumbent will choose policy $y = x$ conditional on x actually being correct. This is equal to $q\sigma_A + (1-q)\sigma_B$ if $x = A$ and $q(1-\sigma_B) + (1-q)(1-\sigma_A)$ if $x = B$. Thus, discovering that the incumbent matched the state with a given policy x is always "good news" about his ability, but the more biased low-ability incumbents are known to be toward that particular policy, the less informative that news is. With these beliefs in hand, we may now calculate the voter's *value of attention* after observing policy x (denoted ϕ^x). Because the voter chooses whether to pay attention after observing the politician's policy choice, best response behavior straightforwardly requires that she always (never) pays attention after policy x whenever $c < (>)\phi^x$.

The value of attention to the voter derives from the possibility that learning the outcome of the incumbent's policy x will improve selection by *changing* her retention decision; if there was no chance that attention would change her vote, then attention would have no decision-relevant value. A crucial implication is that what the voter is looking *for* when she pays attention depends on how she would vote absent that attention, i.e., her posture following x. Specifically, if her posture is favorable ($\mu^x \geq \gamma$) then she pays attention after x in order to find negative information about the incumbent's ability in the form of a policy failure ($\omega \neq x$). Conversely, if her posture is adversarial ($\mu^x \leq \gamma$), then she pays attention after x in order to find positive information about the incumbent's ability in the form of a policy success ($\omega = x$).

Correspondingly, let ϕ_-^x and ϕ_+^x denote the value of negative and positive attention following policy x, respectively. We then have that,

$$\phi_-^x = \delta(1-q) \cdot \Pr(\omega \neq x|y = x)\left(\gamma - \mu_{\neg x}^x\right)$$

$$\phi_+^x = \delta(1-q) \cdot \Pr(\omega = x|y = x);\left(\mu_x^x - \gamma\right).$$

To explain, first observe that the expected net benefit of choosing a high vs. low ability officeholder for the second period is $\delta(1-q)$. The value of negative attention is then this benefit, times the probability $\Pr(\omega \neq x|y = x)$ of uncovering negative evidence, times the difference in probabilities $\gamma - \mu_{\neg x}^x$ that the incumbent and challenger are high ability conditional on that evidence. Similarly, the value of positive attention is $\delta(1-q)$, times the probability $\Pr(\omega = x|y = x)$ of uncovering positive evidence, times the difference in probabilities $\mu_x^x - \gamma$ the incumbent and challenger are high ability conditional on that evidence. Lastly, it is easily verified that $\phi_-^x < (>)\, \phi_+^x$ if and only if the voter has a strictly favorable (adversarial) posture toward the incumbent following x. Thus, the true value of attention following x is simply $\phi^x = \min\{\phi_-^x, \phi_+^x\}$, and the voter best-response is as follows.

Lemma 9 *The voter's strategy is a best response if and only if* $\forall x \in \{A, B\}$*:*

- *Her posture following* x *is strictly favorable (adversarial) when* $\mu^x > (<)\gamma$*;*
- *She always (never) pays attention following policy* x *when the cost of attention* c *is strictly greater than (less than) the value of attention* $\phi^x = \min\left\{\phi_-^x, \phi_+^x\right\}$*;*
- *After paying attention, she never retains an incumbent who mismatched the state, and always (never) retains an incumbent who matched the state when* $\mu_x^x > (<)\gamma$*.*

Attention Absent Pandering

To clarify how rational voter attention works, we last briefly consider equilibrium when there is no ex-ante popular policy, so that both are equally likely to be correct ($\pi = \frac{1}{2}$).

Proposition 10 *If* $\pi = \frac{1}{2}$*, then in equilibrium the incumbent is truthful (*$\sigma_A = 1 > \sigma_B = 0$*). After either policy* $x \in \{a, b\}$ *the voter always retains (replaces) the incumbent absent attention whenever* $\mu > (<)\gamma$*, and always (never) pays attention whenever*

$$c < (>)\, \phi = \delta(1-q) \cdot (\gamma(1-\mu)(1-q) - max\{\gamma - \mu), 0\})$$

Absent any ex-ante difference between the two policies, the voter's treatment of the incumbent in equilibrium cannot depend on his policy choice. Consequently, the incumbent never panders. Despite perfect accountability however, it is still sometimes rational for the voter to pay attention in equilibrium in order to uncover the mistakes of a low-ability incumbent and improve selection.

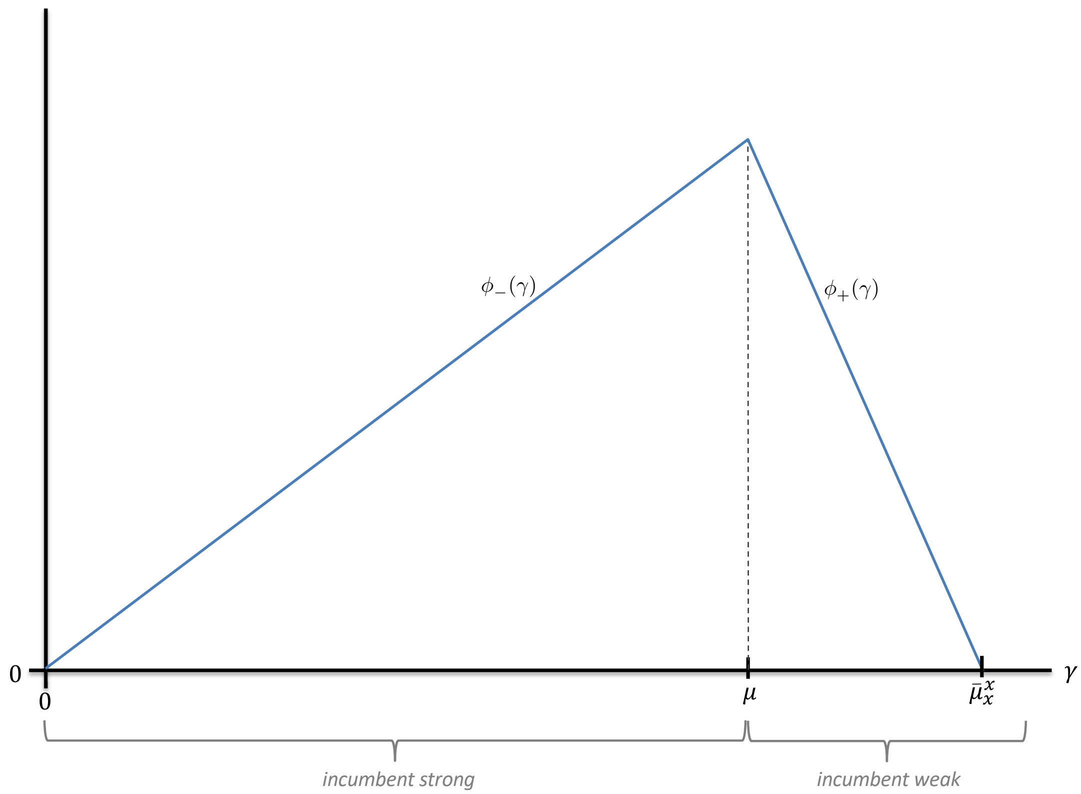


Figure 3.3: Value of attention with symmetric policies.

Figure 3.3 depicts the value of attention as a function of the challenger's reputation γ (holding the incumbent's reputation μ fixed). If the incumbent is initially stronger than the challenger ($\mu > \gamma$), then the value of attention derives from the possibility of discovering that the incumbent's policy mismatched the state, and he is therefore low ability. Consequently, the value of attention ϕ is (locally) increasing in the prior γ that the challenger is high ability. Conversely, if the incumbent is initially weaker than the challenger ($\mu < \gamma$), then the value of attention derives from the possibility that the incumbent's policy matched the state, and he is therefore sufficiently likely to be high ability to justify retention. Consequently, the value of attention ϕ is (locally) decreasing in the prior γ that the challenger is high ability, and becomes 0 when the incumbent is so weak that even matching the state cannot gain him reelection ($\mu_x^x = \frac{\mu}{\mu+(1-\mu)q} \leq \gamma$). Finally, the voter pays the most attention when the race is closest ($\mu = \gamma$), as she has the most to gain from learning about the incumbent's ability.

3.5 Preliminary Results

Recall that there are two ways that a low-ability incumbent might misrepresent his information in equilibrium—(a) by sometimes choosing the ex-ante popular policy A even when his private information indicates that B is correct ($\sigma_B > 0, \sigma_A = 1$), i.e., pandering, or (b) by sometimes choosing the ex-ante unpopular policy B even when his private information indicates that A is correct ($\sigma_B = 0, \sigma_A < 1$), i.e., fake leadership.

In the CHS model with either no attention or symmetric exogenous attention, only pandering can occur in equilibrium. The reason is that the only force distorting the incumbent's policy choice is a strategic incentive to signal competence by choosing the popular policy (see Proposition 8). Asymmetric attention, however, can introduce two additional forces that could, in theory, distort the incumbent's incentives *both* toward pandering or toward fake leadership—an incentive for an initially-strong incumbent to avoid attention, and an incentive for an initially-weak incumbent to seek it. Specifically, if the voter is inclined to retain the incumbent outright after a policy x but pay attention after $\neg x$, then an incentive to avoid attention will bias the incumbent toward x in order to escape the risk that the voter will discover $\neg x$ to have been incorrect and replace him. Alternatively, if the voter is inclined to pay attention after x but replace outright after $\neg x$, then an incentive to seek attention will bias the incumbent toward x in the hopes that the voter will discover x to have been correct and retain him. Indeed, in an extension considered in Canes-Wrone, Herron, Shotts (2001) where the voter exogenously pays more attention after the unpopular policy B ($\rho^A = 0 < \rho^B = 1$), fake leadership can occur in equilibrium when a weak low-ability incumbent chases the attention that the unpopular policy B brings, hoping that this attention will reveal him to have matched the state despite ignoring his private signal.

Our first main result is that when the voter's attention is *endogenous*, fake leadership—driven either by a desire to seek attention or to avoid it—cannot occur in equilibrium. This is true *even though* the voter's attention is generically asymmetric, and does indeed distort the incumbent's policy decisions in equilibrium above and beyond the CHS model.

Proposition 11 *In an equilibrium of the rational attention model, a low-ability incumbent never exercises fake leadership, i.e., chooses policy B after observing signal A.*

It is far from obvious that rational voter attention can induce or exacerbate pandering, but never induce fake leadership. The key insight is that the incumbent's interim reputation μ^x does not alone determine how much attention the voter will rationally pay; rather, it interacts with the voter's interim belief $P(\omega = x|y = x)$ that the chosen policy y is correct in a particular way. Specifically, if the incumbent begins sufficiently weak that the voter prefers to replace her even after the popular policy ($\mu^A < \gamma$), then it is the *popular* policy that will receive more attention, because the only information that will change the voter's decision is discovering that the incumbent chose correctly. Conversely, if the incumbent begins so strong that the voter prefers to retain her even after the unpopular policy ($\mu^B > \gamma$), then it is the *unpopular* policy that will receive more attention, because the only information that will change the voter's decision is discovering that the incumbent chose incorrectly. Consequently, when the incumbent prefers to seek attention (because he begins weak) it is precisely pandering that will draw that attention, while when he prefers to avoid attention (because he begins strong) it is *again* pandering that will deflect that attention.

Leadership and Pandering with Rational Attention

Having established that rational voter attention can only distort the incumbent's incentives toward pandering and never fake leadership, we next more closely examine why and when rational attention can either induce a low-ability incumbent to be truthful who would have otherwise pandered, or visa versa. Henceforth we assume that $\sigma_A = 1$ (a low-ability incumbent always chooses the popular policy when his signal indicates it) and denote σ_B (the probability a low-ability incumbent panders after a signal indicating the unpopular policy) as simply σ.

How Rational Attention Can Induce Leadership

Recall that in the CHS model, a low-ability incumbent will pander if and only if the following two conditions hold: (1) he begins relatively even with the challenger ($\gamma \in (\bar{\mu}^B, \bar{\mu}^A)$) (so that the voter will condition retention on policy choice), and (2) his information is sufficiently poor to make pandering profitable ($q < \hat{q}(\delta, \pi)$). The latter is equivalent to:

$$P(\omega = A|s = B) + \delta q > P(\omega = B|s = B).$$

It is easy to see that these two conditions no longer suffice to induce pandering when the voter can pay attention. For example, if she were to pay attention after both policies, then she would clearly restore the incumbent's incentive to be truthful, since matching the state would then maximize both the incumbent's contemporaneous payoffs and his reelection prospects.

More interestingly, however, attention after *only one* policy may also be sufficient to restore the incumbent's incentive to be truthful. To see this, observe that if the voter were to pay attention after only the popular policy A (and still replace outright after B) the incumbent would only have an incentive to pander when:

$$P(\omega = A|s = B)(1 + \delta q) > P(\omega = B|s = B),$$

since now the incumbent will not always be reelected after choosing A, but only when she succeeds with it. Attention after A thus functions as a "punishment" for choosing the popular policy relative to simply retaining the incumbent outright. Similarly, if the voter were to pay attention after only the unpopular policy B (and still retain outright after A) the incumbent would only have an incentive to pander when

$$P(\omega = A|s = B) + \delta q > P(\omega = B|s = B)(1 + \delta q),$$

since now the incumbent may be reelected even after B if it is actually correct. Attention after B thus functions as a "reward" for choosing the unpopular policy relative to simply replacing the incumbent outright.

It turns out either form of asymmetric attention will be sufficient to restore the incumbent's incentive to be truthful (relative to asymmetric attention) if and only if

$$P(\omega = B|s = B) - P(\omega = A|s = B) \geq \delta q \cdot P(\omega = A|s = B),$$

or if the net policy benefit of following the signal $s = B$ exceeds the net future benefit δq of reelection, times the probability $P(\omega = A|s = B)$ that the signal $s = B$ is wrong. The intuition is simple; under either form of asymmetric attention, pandering will actually yield an electoral benefit only when the incumbent's private signal indicating B is actually wrong. It is next straightforward to show that asymmetric attention of either form will restore a low-ability incumbent's incentive to be truthful if and only if the quality q of his information exceeds a unique threshold $\bar{q} \in (\pi, \hat{q})$ that solves

$$\delta\bar{q} \cdot \pi(1 - \bar{q}) = \bar{q} - \pi \tag{3.2}$$

Using the preceding we now state formal conditions under which rational attention will eliminate pandering. (Recall that $\phi^x(\sigma) = \min\left\{\phi^x_-(\sigma), \phi^x_+(\sigma)\right\}$ denotes the voter's value of attention following policy $x \in \{A, B\}$, where we now make the dependence of these quantities on a low-ability incumbent's pandering probability σ explicit.)

Proposition 12 *Say that a low-ability incumbent receives **high-quality** information if $q \in [\hat{q}, 1]$, **moderate-quality** information if $q \in [\bar{q}, \hat{q})$, and **poor-quality** information if $q \in (\pi, \bar{q})$. When $\gamma \in (\bar{\mu}^B, \bar{\mu}^A)$ and $q < \hat{q}$—so that the incumbent will pander absent attention—rational voter attention will eliminate the incumbent's incentive to pander i.f.f. **either**:*

1. *The voter has a low cost of attention ($c \leq \min\{\phi^A(0), \phi^B(0)\}$) and pays attention after both policies;*
2. *The voter has an intermediate cost of attention and pays attention after one policy, i.e.,*

$$c \in (\min\{\phi^A(0), \phi^B(0)\}, \max\{\phi^A(0), \phi^B(0)\}],$$

***and** a low-ability incumbent receives moderate information ($q \in [\bar{q}, \hat{q})$).*

Figure 3.4 indicates the regions of the parameter space within which rational attention induces a *change* in whether a low-ability incumbent is truthful or panders in equilibrium (note that it does *not* also identify the regions where both models exhibit pandering, but to different degrees). The challenger's reputation γ is on the x-axis, while the voter's cost of attention c is on the y-axis. In the lower white pentagon the voter pays symmetric attention even when believing that low-ability incumbents do not pander in order to catch their mistakes. In equilibrium, this attention induces the incumbent to be truthful regardless of the quality of his information. In the upper two dashed triangles the voter pays asymmetric attention when believing that low-ability incumbents do not pander, but this asymmetric attention only restores a low-ability incumbent's incentive to be truthful when he receives moderate-quality information ($q \in [\bar{q}, \hat{q})$). In the larger left triangle, attention restores leadership by effectively rewarding the incumbent for choosing the unpopular policy B. In the smaller white triangle, attention restores leadership by effectively punishing the incumbent for choosing the popular policy A.

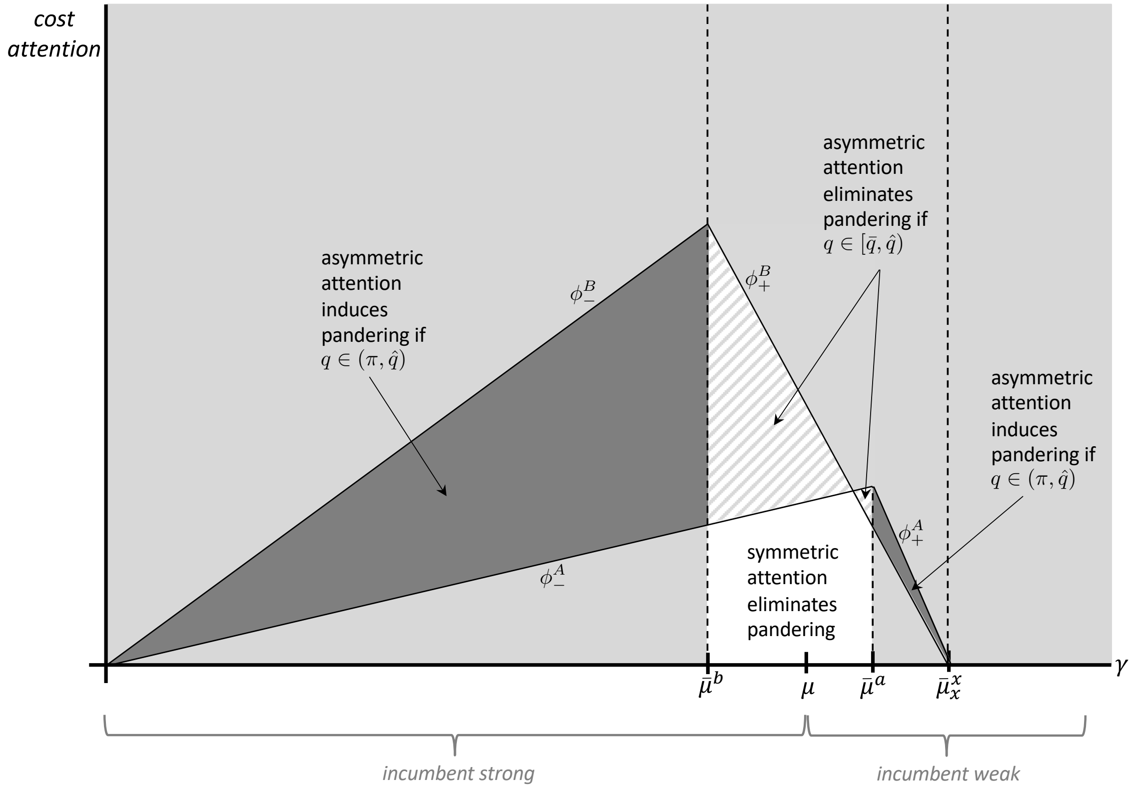


Figure 3.4: Regions where attention eliminates or induces pandering.

How Rational Attention Can Induce Pandering

In the CHS model, a low-ability incumbent is always truthful when he starts out so far ahead of or behind the challenger that the voter will not condition retention on policy, i.e., $\gamma \notin (\bar{\mu}^B, \bar{\mu}^A)$. Clearly, introducing symmetric voter attention will not induce such an incumbent to pander because it simply makes reelection contingent on policy success instead of policy choice. However, we have already shown that asymmetric voter attention reduces, *but does not eliminate*, the incentive to pander. An immediate implication is thus that introducing asymmetric voter attention to an environment in which the incumbent's electoral fate would have otherwise been sealed might induce him to pander. Finally, such asymmetry is a fundamental feature of rational voter attention. The reason is two-fold. First, when the incumbent is sufficiently strong ($\gamma < \bar{\mu}^B$) the voter will want to look for negative evidence after both policies, but be more likely to find it after B. Conversely, when the incumbent is sufficiently weak ($\gamma > \bar{\mu}^A$) the voter will want to look for positive evidence after both policies, but be more likely to find it after A. We thus have the following.

Proposition 13 *When $\gamma \notin (\bar{\mu}^B, \bar{\mu}^A)$—so that the incumbent is truthful in the model* *absent attention—rational attention will induce pandering i.f.f. a low-ability incumbent receives poor information ($q \in [\pi, \bar{q})$)* ***and*** *the voter has an intermediate cost of attention, i.e.,*

$$c \in \left(\min\left\{\phi^A(0), \phi^B(0)\right\}, \max\left\{\phi^A(0), \phi^B(0)\right\}\right].$$

The two darkly shaded triangles in Figure 3.4 depict the regions of the parameter space within which rational attention induces pandering from a low-ability incumbent with poor-quality information. When the incumbent begins sufficiently ahead of the challenger, a voter with an intermediate cost of attention will subject the unpopular policy B to extra scrutiny, inducing a low-ability incumbent to sometimes pander in order to avoid that scrutiny and ensure his reelection. Conversely, when the incumbent begins sufficiently behind the challenger, a voter with an intermediate cost of attention will grant the popular policy A extra attention, inducing a low-ability incumbent to sometimes pander in order to receive that attention and potentially win reelection. Finally, rational voter attention generally tilted is favor of the unpopular policy; it is therefore more likely to lead a strong incumbent to "play it safe" and avoid the unpopular policy than a weak incumbent to "gamble for resurrection" with the popular one.

3.6 Equilibrium Characterization

Having both ruled out fake leadership, and established when and why rational voter attention can both eliminate and induce pandering, we now fully characterize equilibrium. (In the Appendix we show that the equilibrium level of pandering in the rational attention model is generically unique, and henceforth denote this σ_R^*).

Symmetric Attention

We first provide necessary and sufficient conditions for the voter to pay "symmetric" attention in equilibrium – i.e., the same level of attention after either policy. These conditions may be written simply in terms of the equilibrium pandering level σ_N^* of the CHS model.

Lemma 14 *In an equilibrium of the rational attention model, the voter pays the same level of attention after either policy ($\rho^A = \rho^B$) if and only if either:*

- $c < \min\{\phi^A(0), \phi^B(0)\}$*, so that the voter pays full attention after both policies ($\rho^A = \rho^B = 1$) and the incumbent never panders*

- *$c > \max\{\phi^A(\sigma_N^*), \phi^B(\sigma_N^*)\}$, so that the voter never pays attention after either policy ($\rho^A = \rho^B = 0$), and the incumbent panders to the same degree σ_N^* as in the CHS model*

Finally, there exists some $\underline{\gamma} \in (\mu, \bar{\mu}^A)$ at which $\phi^B(0)$ crosses $\phi^A(0)$, and another $\bar{\gamma} \in (\underline{\gamma}, \bar{\mu}^A)$ at which $\phi^B(\sigma_N^*(\gamma))$ crosses $\phi^A(\sigma_N^*(\gamma))$.

Figure 3.5 depicts the two disjoint symmetric attention regions by graphing the values of attention after each policy when the incumbent is believed to be truthful, and when the incumbent is believed to be pandering at level σ_N^*. The darkness of the lines indicates the policy (dark for A, light for B), while the texture indicates expected incumbent behavior (solid for truthful, dashed for pandering). When the cost of attention is below the *minimum* of the values of attention $\phi^A(0) = \min\{\phi_-^A(0), \phi_+^A(0)\}$ and $\phi^B(0) = \min\{\phi_-^B(0), \phi_+^B(0)\}$ *when the voter believes the incumbent to be truthful*, the voter will pay full attention after both policies, which will indeed induce the incumbent to be truthful. These two values of attention cross at a unique challenger reputation $\underline{\gamma} \in (\mu, \bar{\mu}^A)$. Conversely, when the cost of attention is above the *maximum* of the values of attention $\phi^A(\sigma_N^*) = \min\{\phi_-^A(\sigma_N^*), \phi_+^A(\sigma_N^*)\}$ and $\phi^B(\sigma_N^*) = \min\{\phi_-^B(\sigma_N^*), \phi_+^B(\sigma_N^*)\}$ *when the voter believes the incumbent to be pandering at level* $\sigma_N^* \geq 0$, then the voter will never pay attention after either policy, and the incumbent will indeed behave as if attention is impossible. These two values of attention again cross at a unique challenger reputation $\bar{\gamma} \in (\underline{\gamma}, \bar{\mu}^A)$.

Asymmetric Attention

Having identified the conditions under which the voter will pay symmetric vs. asymmetric attention in equilibrium, we next characterize which policy will elicit more attention in equilibrium. For use in this and subsequent propositions, let $\sigma_{x,s}^{x',s'}$ denote the level of pandering that satisfies the equality $\phi_s^x(\sigma_{x,s}^{x',s'}) = \phi_{s'}^{x'}(\sigma_{x,s}^{x',s'})$ where $x \in \{A, B\}$ and $s \in \{-, +\}$; so for example, σ_{A-}^{B+} is the level of pandering that will induce the voter to equally value negative attention after policy A and positive attention after policy B.[8] Further, by Lemma 9 we have that $\mu^x(\sigma_{x-}^{x+}) = \gamma$. In other words, the level of pandering σ_{x-}^{x+} that equates the values of positive and negative attention after a policy x is exactly the level of pandering that makes the voter indifferent over retaining the incumbent after x absent attention. Thus, in the (now obsolete) notation of Section 3.4 characterizing the CHS model, we

[8] In the Appendix we derive these six quantities more precisely and derive a variety of properties.

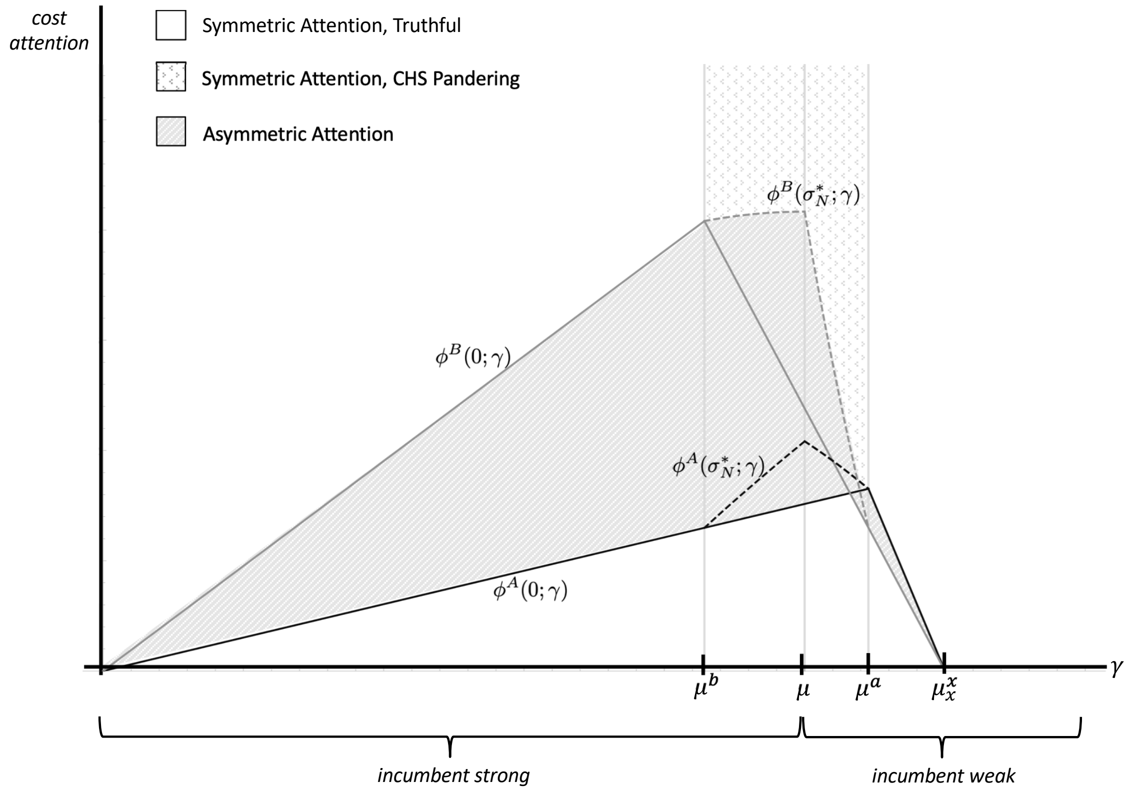


Figure 3.5: Regions with symmetric and asymmetric attention in rational attention model.

have that $\hat{\sigma}_B^x = \sigma_{x-}^{x+}$, and therefore equilibrium pandering in the CHS model is $\sigma_N^* = \min\{\max\{\sigma_{A-}^{A+}, 0\}, \max\{\sigma_{B-}^{B+}, 0\}\}$.

Proposition 15 *Suppose that the voter pays asymmetric attention in equilibrium.*

- *If $\gamma < \underline{\gamma}$ then the voter pays more attention after policy B*
- *If $\gamma > \bar{\gamma}$ then the voter pays more attention after policy A*
- *If $\gamma \in \left[\underline{\gamma}, \bar{\gamma}\right]$ then the voter pays more attention after policy B (A) if*

$$c > (<)\ \phi_+^B\left(\sigma_{A-}^{B+}\right) = \phi_-^A\left(\sigma_{A-}^{B+}\right)$$

Figure 3.6 illustrates the regions of the parameter space within which the voter pays more attention to each policy. A rational voter clearly exhibits a strong attentional bias toward the unpopular policy *B*, which is surprising given that it is the popular policy *A* chosen by "panderers." The intuition is as follows.

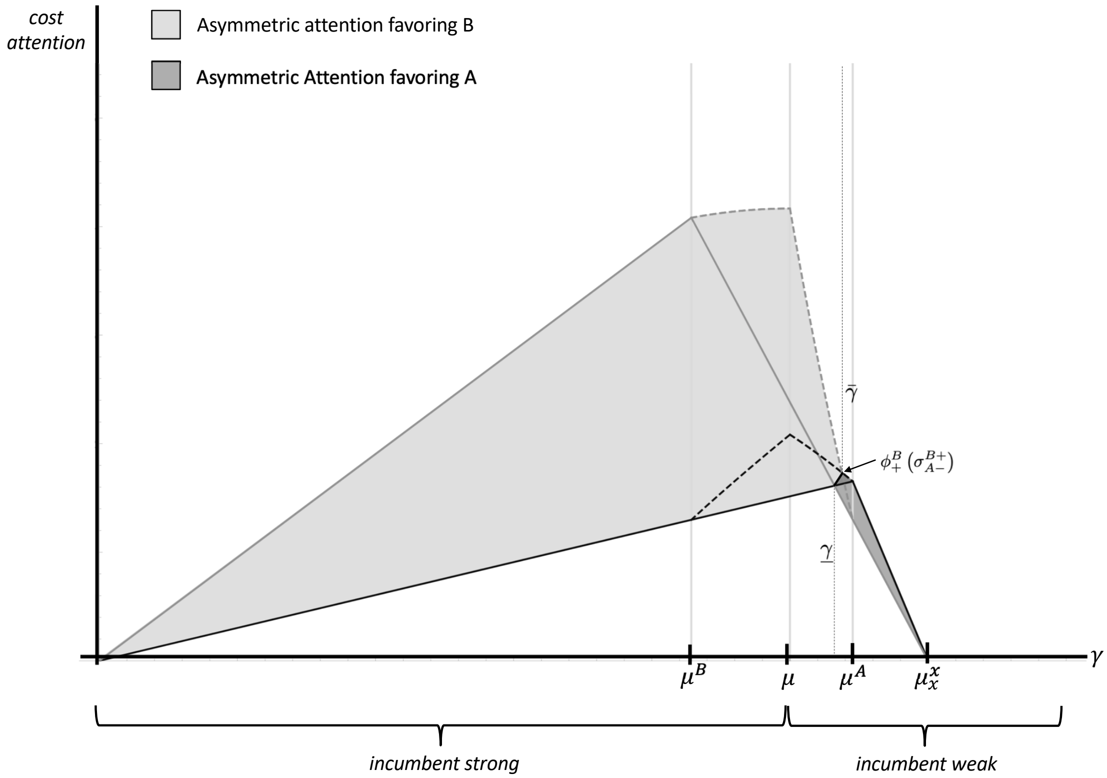


Figure 3.6: Regions where attention is biased toward each policy

The voter's willingness to pay attention to the unpopular policy *B* dominates if the challenger is weak, when it is motivated by the incentive to discover the incumbent made a mistake and should be replaced. Even in the presence of pandering, this incentive is strong given B's initial unpopularity and the strength of the evidence about incumbent ability contained in a visible policy failure. Thus, under these conditions there is a comparatively large range of attention values where the voter is willing to pay attention after policy *B* but not policy *A*. Conversely, the voter's willingness to pay attention to the popular policy *A* dominates if the challenger is strong, when it is motivated by the incentive to discover that the incumbent actually chose correctly and should be retained. However, the presence of a strong challenger makes information about the incumbent less valuable overall; in addition, a policy success is weaker evidence about the incumbent's abilities than a policy failure. Consequently, whenever the voter is *more* willing to pay more attention to policy *A* than *B*, she is also not very willing to pay attention overall. Finally, when the incumbent is relatively even with the challenger, the voter pays attention to find different types of information after each policy—specifically, she pays attention after *A* to "catch" a low-ability incumbent who is pandering, but after *B* to "uncover" a high-ability incumbent who is exercising leadership. However, it is easier to do the latter than the former—a low-ability panderer is so incompetent that he may actually achieve a policy success when he meant to pander, but a high-ability leader will always achieve a policy success when he meant to exercise leadership. Thus, under these conditions the incentive to pay attention after B largely dominates.

Attention Favoring B

We next characterize the exact form of equilibrium when the voter pays more attention to B, beginning with case of a low-ability incumbent who receives moderate-quality information.

Proposition 16 *Suppose that the voter pays more attention to policy B in equilibrium, and a low-ability incumbent receives moderate-quality information ($q \in [\bar{q}, \hat{q}]$). Then the voter always retains the incumbent outright after policy A ($\nu^A = 1 > \rho^A = 0$).*

- *If the voter is willing to pay attention to B given truthfulness ($c < \phi^B(0)$), then then in equilibrium the incumbent is truthful, and the voter pays full attention to B ($\rho^B = 1$).*
- *If the voter is unwilling to pay attention to B given truthfulness ($c > \phi^B(0)$), then in equilibrium the incumbent panders ($\sigma_R^* > 0$), but strictly less than in the CHS model ($\sigma_R^* < \sigma_N^*$), and the voter pays some attention to B ($0 = \nu^B < \rho^B < 1$).*

Equilibrium when a low-ability incumbent receives moderate-quality information is depicted in Figure 3.7; the voter pays more attention to B in the lightly shaded area. As described in Section 3.5, the incumbent is truthful when the voter is willing to pay attention to B given the expectation of truthfulness. When she is not, the unique equilibrium involves partial attention after B ($\rho^B \in (0, 1) > \rho^A = 0$)—just enough to make a low-ability incumbent indifferent to pandering. The incumbent in turn panders, but just enough to make the voter indifferent to paying attention after B.

We next turn to the more complex case of a low-ability incumbent who receives poor-quality information ($q \in (\pi, \bar{q})$). As discussed in Section 3.5, when a low-ability incumbent receives poor information, asymmetric attention can create or exacerbate the incentive to pander. As a consequence, rational attention can have a

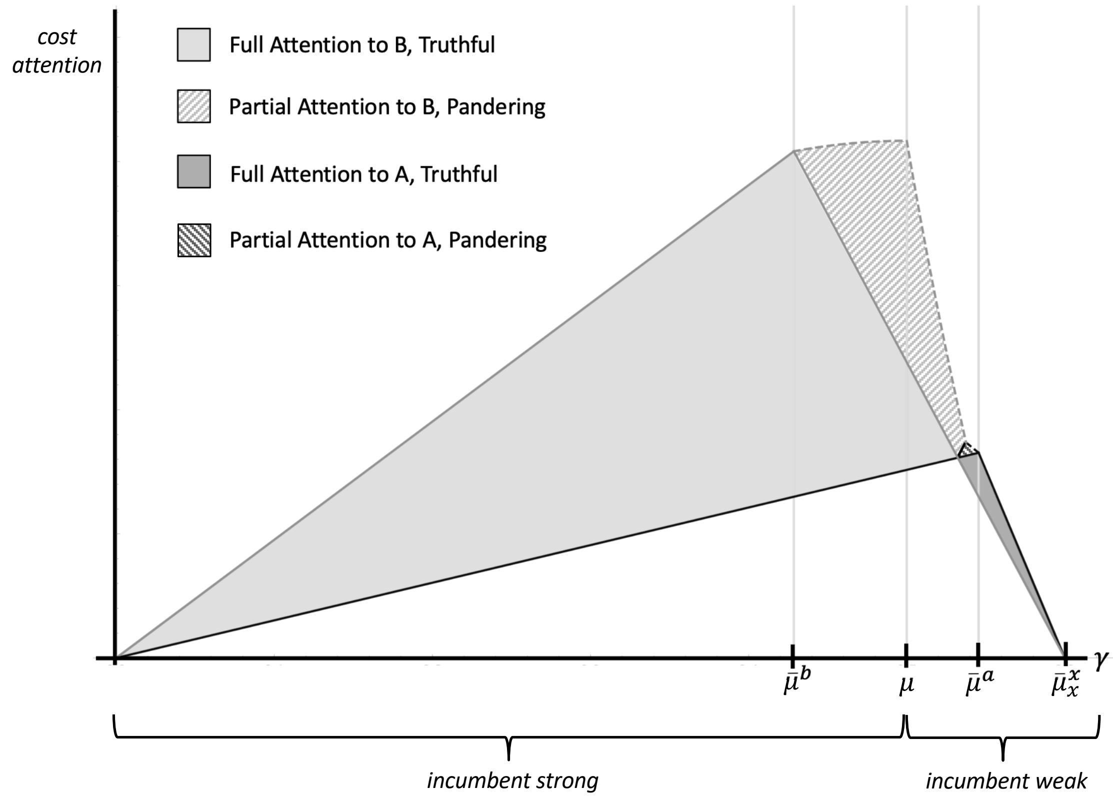


Figure 3.7: Asymmetric attention equilibria with moderate information

variety of effects on the incumbent's equilibrium behavior. It can decrease (but not eliminate) pandering that would have otherwise occurred absent attention, induce pandering that would not have occurred absent attention, and even worsen pandering that would have already occurred absent attention.

Proposition 17 *Suppose that the voter pays more attention to policy B in equilibrium and a low-ability incumbent receives poor-quality information ($q \in (\pi, \bar{q})$) Then the incumbent always panders ($\sigma_R^* > 0$) to avoid the attention that the unpopular policy brings.*

- *If $c < \min\{\phi_-^A(\sigma_{A-}^{B-}), \phi_-^A(\sigma_{A-}^{A+})\}$, then the voter always pays attention after policy B ($\rho^B = 1$) and sometimes after policy A ($\rho^A \in (0,1)$ and $\nu^A = 1$).*

- *If $c > \max\{\phi_-^A(\sigma_{A-}^{B-}), \phi_-^B(\sigma_{A-}^{A+})\}$, then the voter sometimes pays attention after policy B ($\rho^B \in (0,1)$ and $\nu^B = 1$) and never after policy A ($\rho^A = 0$ and $\nu^A = 1$).*

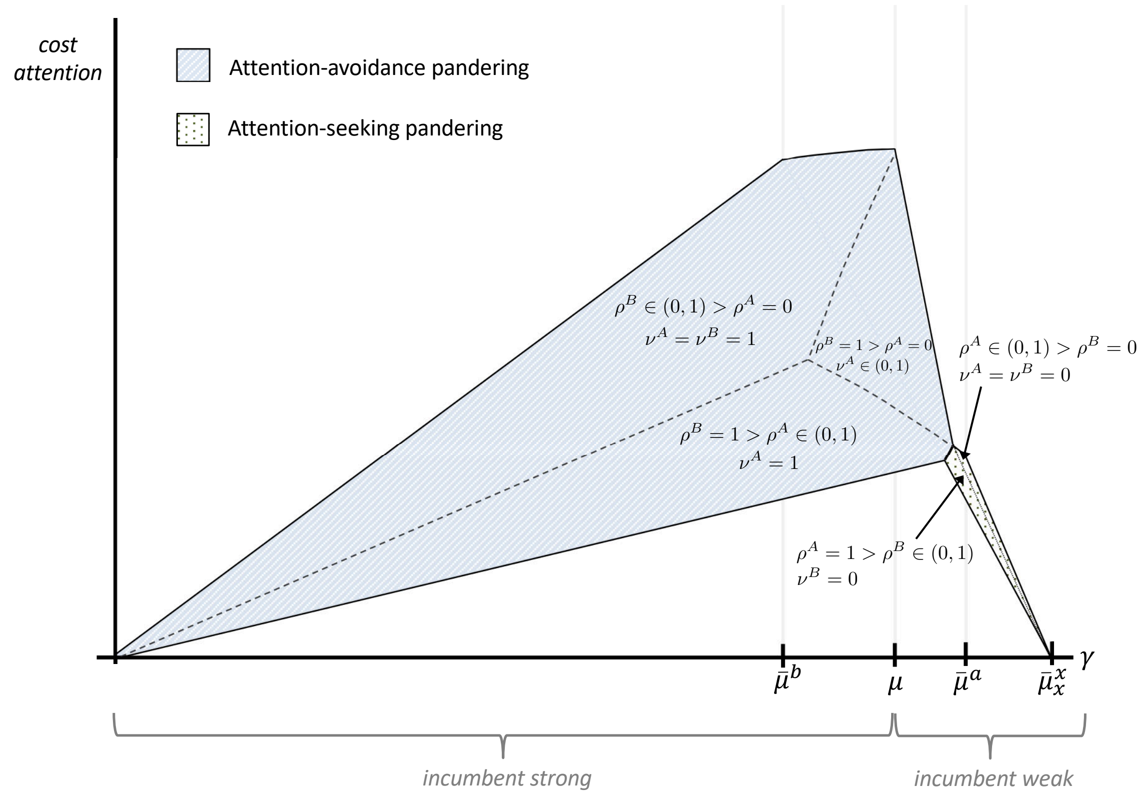


Figure 3.8: Asymmetric attention equilibria with poor information

- *If* $c \in [\phi_-^A(\sigma_{A-}^{A+}), \phi_-^B(\sigma_{A-}^{A+})]$*, then the voter always pays attention after policy* B *(*$\rho^B = 1$*), and sometimes retains but never pays attention after policy* A *(*$\nu^A \in (0,1)$ *and* $\rho^A = 0$*).*

Equilibria when a low-ability incumbent receives poor-quality information are depicted in Figure 3.8. The voter pays more attention to B in the lightly shaded area, and within this area there are three types of equilibria. In the first, the voter pays a "high" amount of attention – always after policy B and sometimes after policy A. Moreover, the incumbent panders *more* as attention becomes costlier, because this exacerbates the voter's propensity to pay different levels of attention to the two policies. For sufficiently high attention costs the incumbent actually panders strictly more than in the CHS model ($\sigma_R^* > \sigma_N^*$), and choosing the popular policy A perversely becomes an unfavorable signal about the incumbent's ability. In the second, the voter pays a "medium" amount of attention—always after policy B and never after policy A—but sometimes replaces the incumbent outright after policy A. The incumbent's pandering is unaffected by the cost of attention, but is strictly worse than in the CHS model if incumbent is initially strong ($\mu > \gamma$). In the third, the voter pays a "low" amount of attention—sometimes after policy B and never after policy A. Here the incumbent panders less as attention becomes costlier, because this again diminishes the voter's propensity to pay different levels of attention after the two policies. Eventually, attention becomes too costly to be paid after either policy, and equilibrium becomes identical to the CHS model.

Attention Favoring A

We last characterize equilibrium when the voter's attention is biased toward A, again beginning with the case of a low-ability incumbent who receives moderate-quality information.

Proposition 18 *Suppose that the voter pays more attention to policy A in equilibrium, and a low-ability incumbent receives moderate-quality information ($q \in [\bar{q}, \hat{q}]$). Then the voter always replaces the incumbent outright after policy B ($\nu^B = \rho^B = 0$).*

- *If the voter is willing to pay attention to A given truthfulness ($c < \phi^A(0)$), then in equilibrium the incumbent is truthful, and the voter pays full attention after A ($\rho^A = 1$).*
- *If the voter is unwilling to pay attention to A given truthfulness ($c > \phi^A(0)$), then in equilibrium the incumbent panders ($0 < \sigma_R^*$) but strictly less than in the CHS model ($\sigma_R^* < \sigma_N^*$), and the voter pays some attention after A ($0 = \nu^A < \rho^A < 1$).*

The voter pays more attention to A in the darkly shaded area of Figure 3.7; the structure of equilibrium closely resembles the case of a voter who pays more attention to policy B.

Conversely, when a low-ability incumbent receives low-quality information, equilibrium with attention favoring A is as follows.

Proposition 19 *Suppose that the voter pays more attention to policy A in equilibrium and a low-ability incumbent receives poor-quality information ($q \in (\pi, \bar{q})$. Then the incumbent always panders ($\sigma_R^* > 0$) to seek the attention that the popular policy brings.*

- *If* $c < \phi_+^A(\sigma_{A+}^{B+})$*, then the voter always pays attention after policy A (*$\rho^A = 1$*) and sometimes after policy B (*$\rho^B \in (0,1)$ *and* $\nu^B = 0$*).*
- *If* $c > \phi_+^A(\sigma_{A+}^{B+})$*, then the voter sometimes pays attention after policy A (*$\rho^A \in (0,1)$ *and* $\nu^A = 0$*) and never after policy B (*$\rho^B = \nu^B = 0$*).*

The voter pays more attention to A in the darkly shaded area of Figure 3.8. The structure of equilibrium again resembles the case of a voter who pays more attention to policy B in that there are also high and low-attention regions, and within these regions a higher cost of attention affects pandering in the same manner as previously discussed. However, there is no equilibrium in which the voter always pays attention after A and never after B (and therefore sometimes retains outright after B). Such an equilibrium would require the incumbent to be weak, the voter to be willing to retain her after B, but to also be more willing to pay attention after A. However, the voter will only retain a weak incumbent who chooses B if she believes pandering to be very severe—so severe that she will also become more optimistic about B's prospects for success, and more willing to pay attention to it.

3.7 Voter Welfare

Since rational voter attention sometimes comes at the expense of electoral accountability, we conclude our analysis by comparing the voter's equilibrium utility in the rational attention model and the CHS model absent attention. This comparison can be interpreted in two ways. First, it could represent the difference between a setting in which the voter's attention costs are low enough that the ability to pay attention meaningfully impacts her behavior, and one in which those costs are so prohibitive that it is *as if* attention is impossible. Second, it could represent the comparison between a setting in which the information sources that the voter can pay attention to actually contain useful information about incumbent performance, and one in which those information sources are either absent or uninformative.

To ease the exposition we first provide a simplified characterization of the voter's utility difference between the two models that exploits properties of equilibrium.

Lemma 20 *The voter's equilibrium utility difference between the rational attention and CHS models may be written as*

$$\begin{aligned} U_V^R - U_V^N \;=\; & \Pr(y = A) \cdot \max\left\{\phi_s^A - c, 0\right\} + \Pr(y = B) \cdot \max\left\{\phi_s^B - c, 0\right\} \\ & - (1-\mu)(q-\pi)\left(\sigma_R^* - \sigma_N^*\right), \quad \textit{where } s = - \textit{ if } \gamma \le \mu \textit{ and } s = + \textit{ if } \gamma \ge \mu. \end{aligned}$$

All quantities are evaluated with respect to σ_R^* unless explicitly indicated otherwise.

The voter's utility difference between the two models consists of two components. The first is the second-period selection benefit of being able to learn the policy outcome and make a better-informed retention decision. This benefit in turn consists of the unconditional probability $\Pr(y = x)$ that each policy will be chosen, times the value of attention ϕ_s^x conditional on that policy being chosen less the cost c of attention.[9] The second component is the first period accountability cost of increased pandering $\sigma_R^* - \sigma_N^*$ (which is actually a benefit if attention also reduces pandering).

With this characterization in hand, it is simple to state the welfare consequences of attention when a low-ability incumbent receives moderate information ($q \in [\bar{q}, \hat{q}]$)).

Proposition 21 *When a low-ability incumbent receives moderate-quality information, the voter is always weakly better off in the rational attention model, and strictly better off i.f.f. she pays some attention in equilibrium ($\exists\, x \in \{A, B\}$ s.t. $\rho^x > 0$).*

As previously discussed, when a low-ability incumbent receives moderate-quality information, even asymmetric attention will restore his incentive to be truthful. Consequently, in equilibrium the ability to pay attention always weakly benefits the voter, and strictly benefits her when she actually pays some attention in equilibrium (since attention is always associated with strictly better accountability and sometimes strictly better selection as well).

In the case of an incumbent who receives poor-quality information ($q \in [\pi, \bar{q}]$), there is a potential tradeoff between accountability and selection as follows.

Proposition 22 *When a low-ability incumbent receives poor-quality information, there is a unique cost cutpoint $\hat{c}(\gamma)$ such that that the voter is strictly worse off in the rational attention model i.f.f. $c \in (\hat{c}(\gamma), \max\{\phi^A(\sigma_N^*), \phi^B(\sigma_N^*)\})$.*

[9] Worth noting is that the selection benefit is calculated *as if* the voter will always have a favorable posture (adversarial) posture toward an initially strong (weak) incumbent, even if these are not her equilibrium postures in the rational attention model. The reason is that they *are* her equilibrium postures in the CHS model. To clarify the implications of this subtlety, consider an initially-strong incumbent who panders in both models (so $\gamma \in [\bar{\mu}^B, \mu]$), but to a lesser degree in the rational attention model (so $\mu^B < \gamma$ and $\rho^B = 1$). Then in equilibrium the voter actually looks for *positive* information after policy B, so the true interim benefit of attention is ϕ_+^B. The expression in Lemma 20 then embeds an additional selection benefit $\phi_-^B - \phi_+^B$ that the voter *would* enjoy from reduced pandering were she to deviate to paying no attention after B.

- *If* $\gamma < \mu$, *then* $\hat{c}(\gamma) \in (\phi_-^A(0), \max\{\phi_-^B(\sigma_{A-}^{B-}), \phi_-^B(\sigma_{A+}^{A-})\})$
- *If* $\gamma \in (\bar{\gamma}, \bar{\mu}_x^x)$, *then* $\hat{c}(\gamma) \in (\phi_+^B(0), \phi_+^A(\sigma_{A+}^{B+}))$
- *Otherwise* $\hat{c}(\gamma) = \max\{\phi^A(\sigma_N^*), \phi^B(\sigma_N^*)\}$

Figure 3.9 recreates Figure 3.8, but explicitly indicates the two regions within which rational attention strictly harms voter welfare. Although the notation of Proposition 22 is cumbersome, the interpretation is straightforward. First, it is immediate that within the two "attention-avoidance" and "attention-seeking" regions where the voter's attention is also low ($1 > \rho^x > 0 = \rho^{\neg x}$), the voter must be strictly worse off in the rational attention model. The reason is that within these regions, the voter enjoys no selection benefit from paying attention (since she either strictly or weakly prefers not to), but suffers a strictly positive accountability cost (since the incumbent panders strictly more than in the CHS model). Second, moving away from the boundaries of these regions—by either lowering the cost of attention c or shrinking the difference in candidate reputations $|\gamma - \mu|$—must strictly increase voter welfare through some combination of better selection and better accountability. Finally, along the boundaries that separate the asymmetric attention and full attention regions, the voter must be strictly better off in the rational attention model, since she enjoys a strictly positive selection benefit from attention but no accountability cost (since the incumbent is truthful).

3.8 Conclusion

In this paper we consider a variant of the canonical political agency model of **?** in which the voter must pay an attention cost to learn about the consequences of the incumbent's policy. Our model is intended to study political accountability in environments where it is not information about incumbent performance that is scarce, but rather the voters' attention in consuming and processing such information.

Our key findings are as follows. First, rational voter attention will be asymmetric across different policy alternatives when the voter's cost of attention is moderate. The reason is that the voter's willingness to pay attention is determined by his belief that such attention will uncover information that reverses his voting intention based on the observed policy alone, and these beliefs generically differ across the two policy alternatives if one is initially more-popular. Specifically, if the voter's current voting intention is to retain the incumbent, then she will only pay attention to uncover a failure that would justify replacing him. Alternatively, if her current voting intention

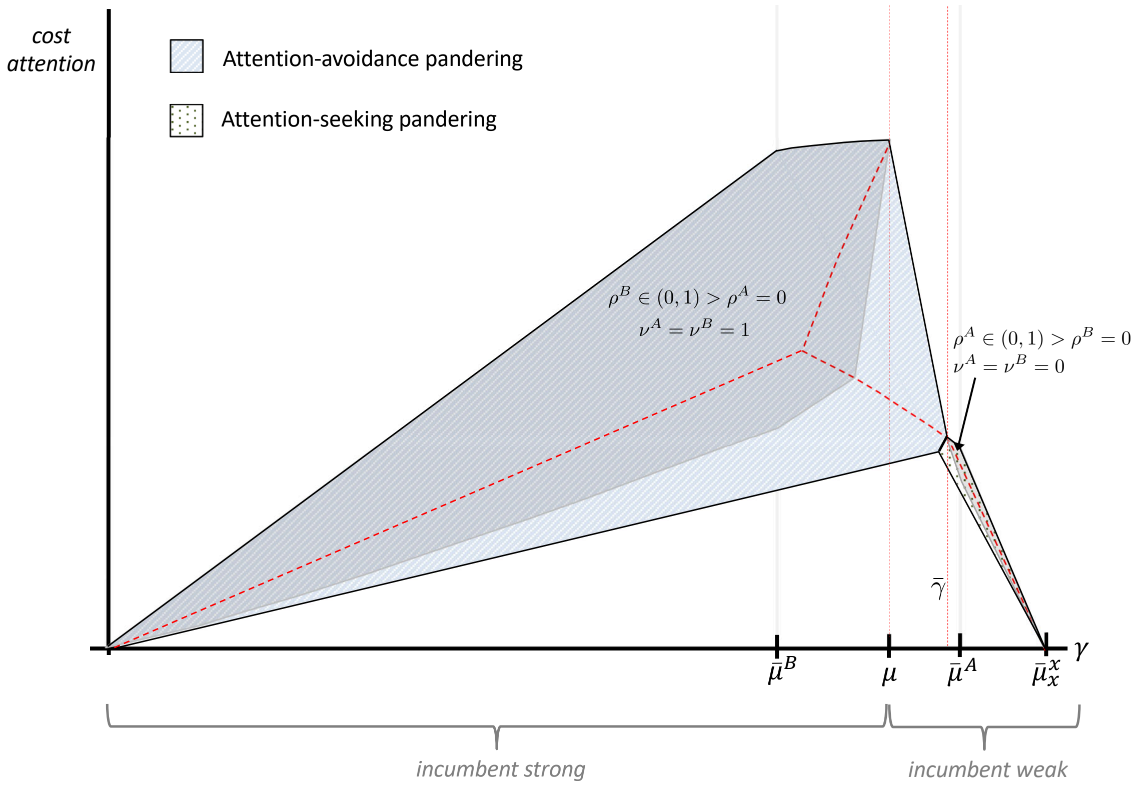


Figure 3.9: Where rational voter attention strictly worsens voter welfare.

is to replace the incumbent, then she will only pay attention to uncover a success that would justify retaining him. These effects make a rational voter generally more willing to pay attention after an unpopular policy than a popular one. The reason is that the prospects for uncovering a "leader" who chose the unpopular policy are better than the prospects for uncovering a "panderer" who chose the popular policy.

Second, rational attention can improve electoral accountability—by rewarding the incumbent with attention for choosing the unpopular policy, punishing the incumbent with attention for choosing the popular policy, or both. However, it can also harm electoral accountability, by giving an ex-ante strong incumbent an incentive to choose the policy that evades attention, or by giving an ex-ante weak-incumbent an incentive to choose the policy that draws it. Both of these effects can worsen pandering relative to a setting in which the voter cannot learn about policy consequences at all, and these effects can be sufficiently strong to harm voter welfare despite the improvement in selection that attention brings. However, they cannot induce the phenomenon of fake leadership (choosing the ex-ante unpopular policy to draw or evade attention attention) as uncovered in the original **?** model with exogenous information revelation.

Our positive results about rational voter attention yield several testable empirical implications. First, if "paying attention" in the model is interpreted as an increase in the consumption of political media, then a clear implication is that it is *unpopular* policy choices that will generally drive political media consumption. Moreover, the model suggests that this increased media consumption will be specifically motivated by the desire to "find reasons" that the incumbent's policy choice was not as misguided as it appeared to be. Second, the model adds to the set of conditions under which incumbent politicians can be expected to follow public opinion. The classic **?** model highlights the relationship between pandering and electoral competitiveness. To this we add a relationship between pandering and the sensitivity of voters' consumption of political information. If voters' consumption of political information is largely insensitive to policy choice—either because they always consume a great deal of it or little to none—then incumbents will be more likely to follow their own guidance. However, if voters' consumption of political information is very sensitive to policy choice—because doing so is somewhat but not prohibitively costly—then incumbents will be more likely to follow popular opinion to shape those consumption choices.

Finally, with respect to implications about voter welfare the model adds to a small but growing literature that identifies reasons why improving voters' informational environment—either by improving the accuracy of information sources or by lowering the costs to voters of acquiring political information—may be a double edged sword for both politicians and voters (e.g. **?**). In our model, both may be harmed by such improvements if they do not overcome, or even exacerbate, the propensity of rational voters to apply different levels of scrutiny to different policy choices.

Our model also suggests several avenues for future research; we comment on two in particular. First, a now large literature considers voters' choices over the consumption of biased media (**?????**). Our model can be easily extended in this direction by allowing the voter to choose between two binary noisy signals of the incumbent's policy outcome when she pays attention—one that is biased in favor of the incumbent, and another that is biased against. Since a key feature of the model's logic is that the voter is looking *for* something in particular when she pays attention, such an extension could shed light on how voters' choice of news media is influenced by both an incumbent's actual policy choices and by his competitive environment. Second, an existing literature examines what sort of media landscape best promotes well-informed voting and political accountability (**???**). An extension of our model could shed further light on this question by having the voter choose whether to pay attention to a noisy (rather than perfect) binary signal about incumbent performance, and analyzing features of the conditional probability distribution over this signal that improve or maximize voter welfare once her rational attention decisions are taken into consideration. In particular, the logic of the model suggests that democratic accountability may be improved if the media environment is biased in a manner that somehow counterbalances the voter's natural propensity to apply different levels of scrutiny to different policy choices. We hope to explore these and other avenues in future work.

Chapter 4

FORMAL REQUIREMENTS ON COSTLY INFORMATION

4.1 Introduction

Motivation

The U.S. government heavily relies on federal advisory committees that provide recommendations for different government agencies. The Federal Advisory Committee Act (FACA) limits the power of these committees to administer advice, which may or may not be used by a specific government agency or the executive branch. Even though the committees do not make any policy decisions, their overall level of influence is high. There is strong empirical evidence suggesting that these committees often make biased recommendations. As a consequence, the literature mainly assumes their preferences are biased, and the majority of the work on this topic explores agency problem between government agencies and these biased committees.[1]

We take research on federal advisory committees a step further in this study, formulating a formal theoretic model that explains where the bias in elected policies comes from, even when the final policy decision is made by an unbiased committee. We also show how procedural requirements can be imposed by the committee itself (rather than upon them) in order to improve the fairness of the policy choice and increase overall welfare. Our approach suggests that bias in policy decisions stems from another agency problem within the committee's working procedures that cannot be resolved with standard "auditing" or procedural requirements often used in the bureaucratic oversight literature. We illustrate how the optimal mechanism based on formal requirements might work.

We construct a model with an unbiased principal (committee) who makes a final policy choice optimal for the unobserved binary state of the world. The principal relies on a proposal consisting of private/soft and public/hard information from the biased agent (the proposer). The agent's preferences are biased toward electing the highest possible policy. Private information is costless for the agent and unverifiable for the principal. Public information requires some costly effort from the agent, this information is publicly observable and unfalsifiable. We analyze the equilibria

[1]There is a small empirical literature on the effects of administrative procedures on agency performance (**????**).

of two different models: (1) the agent can freely choose the amount of costly public information supporting her proposal; and (2) the principal imposes formal requirements on the amount of public information conditional on the private report from the agent. Since the agent has private information and her preferences are biased, our problem is a mix of the literature on delegated information acquisition and the persuasion literature. The idea is that if the principal does not control the amount of public information required, the biased agent can manipulate it in her favor to induce higher policy outcomes.

In order to better understand our motivation, we apply our model to Medicare pricing of medical services. The committee of physicians, also known as the RVS Update Committee (RUC), makes a final decision about the price of a specific medical service based on the proposal made by the group of specialists/doctors (or just one doctor).[2] We assume that the committee wants to implement fair prices for medical services. In the real world, proposers have to provide a subjective opinion based on their expertise (soft information) as well as the results of a survey (hard information) to support their price evaluation. The fact that proposers usually perform the medical service under consideration makes their preferences biased towards higher prices. In reality, the proposer is free to choose the number of people surveyed and this number varies across the proposals. Therefore, the proposer could in principle manipulate the number of people surveyed to achieve higher prices.

Our results support this intuition. When agents freely choose the amount of public information, for a relatively costless public signal they do not reveal their private information. Instead, in the best possible equilibrium for the principal, the agent always acquires the maximum public information (not revealing her private signal) and the principal makes a decision just based on this information. However, for relatively costly public signals, separation is possible. In equilibrium, the agent observing the low private signal acquires no public information and her proposal consists of only her report of the private signal. Since the agent with a low private signal expects the public signal to also point towards the lower policy choice, she decides not to acquire any in order to avoid such an undesirable outcome. The preferences of the agent observing the high private signal coincide with those of the principal. In equilibrium, such an agent acquires some small amount of public signal, just enough to differentiate herself from the low agent and make her report of private signal believable.

[2]Medicare uses a physician payment system based on the resource-based relative value scale (RBRVS).

In contrast, when the principal can impose formal requirements on the amount of public signal, the amount of public signal serves two different purposes: (1) more public signal increases the precision of the policy choice; and (2) the number of surveys can be used as a tool for extracting the private signal from the agent, if the requirements of the principal are conditioned on the "soft information" reported by the proposer. Our analysis shows that if public information is too cheap, it fails to serve the second purpose. However, for high enough costs, the optimal mechanism requires maximum evidence (in terms of public signal) to support the proposal with high private report. Since the low private report goes against the agent's preferences, such a proposal must be accompanied by less public evidence. This mechanism incentivizes the agent to reveal her low private signal by relieving her of some of the burden of acquiring a costly public signal.

Welfare analysis shows where the benefits of formal requirements come from. For high enough costs, separation (truthfully revealing the agent's private information) is possible in both models. However, the motivation and mechanism behind separation are different based on who takes control over the possible amount of the public signal. When the decision is in the agent's hand, the agent with a high private signal acquires just enough public signal to differentiate herself from the low agent.[3] As the cost of the public signal increases, it gets easier for the high type to make her private report believable and, therefore, she collects even less public signal, decreasing the principal's welfare. In contrast, if the principal puts formal requirements on the public signal, then in separating equilibrium, the costly public signal is used as a mechanism to encourage the low type to reveal her private signal by lifting some of the burden of public proof. When the public signal is costly, the formal requirement is a more effective tool in the principal's hand for separation and he can compromise less on the public signal from the low type. Overall, when the principal imposes formal requirements, his equilibrium expected welfare is increasing in cost of information.

Another interesting finding is that agents sometimes separate when the principal would prefer to disregard their private reports for the sake of a better public signal. If the private signal is not too valuable and the cost of the public signal is intermediate, in absence of formal requirements, the agent would truthfully reveal her own type and exert little effort in public signal. Meanwhile, for the same parameters, the principal does not find the separation worth forgoing some amount of public signal. When

[3]We refer to the private signal of the agent as her type.

principal imposes formal requirements, he would rather require maximum public signal from everyone and make a policy decision solely based on the maximum public information.

To summarize, there is empirical evidence that government advisory committees generate biased policy decisions. So far, the literature has primarily focused on assessing the influence of various procedural restrictions imposed by government agencies on these advisory committees (such as transparency, providing feedback on committee recommendations, etc.), assuming that their preferences are biased. Despite their overall influence, however, little theoretical research has been done exploring how these committees work and where the bias in their decisions comes from. Our paper points out the apparent flow in the operating procedures of the advisory committees (allowing the proposer to choose the amount of public signal), resulting in biased decisions even with unbiased principal. We explore the optimal mechanism for procedural requirements within the advisory committee and show how it improves the final policy choice and the overall welfare.

The setup in this chapter is not uncommon in many principal-agent relationships. When the agent has soft private information, usually there is a possibility to require costly verifiable information as well (such as additional market research or cost-benefit analysis, etc.). Therefore, even though the main application of our model is to advisory committees, it can be applied to any principal agent problem with the possibility of costly (for the agent) verifiable public information.

Related Literature

Our paper relates to a large literature on agency problems in political science, especially in bureaucratic oversight. The types of bureaucratic control can be divided into two major categories (**?**): ex-post oversight with direct auditing of the bureaucrat's actions (**?**) or indirect control based on fire alarm system from interest groups (**?**), and ex-ante prevention using different procedural requirements (**??**). The latter is the focus of this chapter. Proponents of the positive theory of administrative procedure have demonstrated that employing procedural requirements for legislative oversight has various benefits. On the other hand, supporters of the "ossification thesis" (**???????**) believe that procedural constraints obstruct the efficient operation of many agencies. The majority of the work on this topic focuses on the interaction between elected officials (such as the president or Congress) and government agencies whose operations show that their preferences are biased. Instead, we concentrate on specifics of the decision making process of the advisory committee to explore the agency problem within, and use verifiable costly information to partially resolve it.

Since the principal in our model uses a costly verifiable public signal to check the soft/private reports of the agent, our paper also relates to the literature on costly verification (**???**) and more closely **?** who study the general delegation framework with costly but feasible verification. The main difference is that "auditing" in our model is a continuous decision since the principal has to choose the effort level $e \in [0, 1]$ and the outcome/precision of the audit is not perfect but depends on this effort level. Moreover, in our case, the "cost of the audit" is completely born by the agent who has state independent preferences.

The main model in this study depicts a situation in which the principal controls the agent's effort level in acquiring the public signal but is unable to commit to an arbitrary decision rule. This model can be classified as a mechanism design problem with limited commitment. The initial motivation of the paper originates from the literature on persuasion with a biased and informed agent. **?** find the necessary and sufficient conditions for the existence of a signal that strictly benefits the sender. In their example, under specific conditions, an agent could construct a truthful signal that would bring her a favorable outcome. Similarly, if there are no formal requirements from the principal, the proposer who always favors higher Medicare prices might be able to construct a signal that would be beneficial for her (although she is now constrained by the evidence from surveys).

More relevant to our example are the findings of **?** about biased election organizations. According to their model, if voters are uncertain about which state is correct but they do not know the size of the electorate, the organizer can ensure a favorable outcome with high probability. The authors show that when the number of participants across states is exogenous but dependent on the state, a biased equilibrium can be sustained in large elections. In their model, the organizer should know the state in advance. This assumption is applicable to our case assuming that the proposer has a good sense of fair prices. The findings of this paper suggest that, even without the assumption that the incentive compatibility of voters is automatically satisfied, the proposer could manipulate the number of people surveyed and guarantee that the outcome of the survey supports her claim about high prices with high probability. Consequently, prices of medical services will be biased up if the committee bases its decision on the results of this survey.

As already mentioned, **?** find empirical support for this hypothesis. The authors also investigate how bias in prices increases with the affiliation of the committee. The difference in our case is that we assume there is no affiliation since the proposer is biased but the committee always wants fair prices. As we will briefly discuss in the last section, introducing partial bias for the committee could be one of the possible extensions of our model.

Findings in the literature about persuasion suggest that the current mechanism used by the RUC, not regulating the number of people surveyed, could result in higher prices. In this paper we show how the RUC can use formal requirements for the amount of hard information gathered to ensure more equitable prices on medical services. A similar question has been investigated by **?**. The authors undertake a game theoretic approach to compare outcomes of three different setups, relaxing the full commitment assumption in different ways. The authors do not use a mechanism design approach, and unlike our case, their model does not assume preexisting private information of the agent.

This paper is also related to the literature studying task delegation both in continuous and discrete time structures. One significant difference, however, is that we do not model sequential search. The player who has the authority decides an effort level to acquire public signal (which can be thought of as the amount of public signal gathered). Moreover, the majority of the literature about delegation focuses on the moral hazard problem, which is not relevant in our case since the effort level exerted is public information and the agent is not able to falsify or hide the public signal. The effort level can affect the utility of the principal directly through the output (**??**) or indirectly through the informativeness of the signal (**?**). In this model, we undertake the second approach, linking the costly effort to the informativeness of the public signal. An important feature of the model is the pre-existing private information of the agent, which is valuable for the principal.

The rest of this chapter is organized in the following way: Section 2 formally sets up the model for the problem discussed so far. Section 3 does some preliminary analysis for the optimal policy choice. Section 4 shows pooling and separating equilibria for the biased agent who has the authority to choose her own effort level but the principal makes the final decision. This case is the closest to the real-world procedure of deciding prices of medical services by the RUC. Section 5 analyzes the optimal mechanism for the main model in which the principal formally requires a certain level of effort after each report of soft information from the agent but cannot commit to a decision rule. Section 6 is devoted to the welfare comparison of the two models and further explains the mechanism through which the formal requirements increase welfare. Formal proofs as well as the exact values for cutoffs and bounds for the regions of the parameter space are provided in the appendix.

4.2 The Model

We consider a one-period model with two players: the principal (RUC) and the agent (specialist in an area who makes a proposal). In order to avoid pronoun confusion, we refer to the principal as "he" and the agent as "she." Nature draws a binary state $\omega \in \{0, 1\}$ which determines the desirable policy choice for the principal.[4] The principal has to make a final policy decision $x \in [0, 1]$. With respect to the RUC example, a state of the world can be interpreted as the true price of a certain medical procedure being either high (1) or low (0).

Information: The agent receives a private signal s about the state of the world. The precision of this signal is θ, meaning:

$$P(s = 1|\omega = 1) = P(s = 0|\omega = 0) = \theta,$$

where $\theta \in [0.5, 1]$ can be interpreted as the competence level of the agent. Even without further investigation, based on her professional opinion, the proposer has some idea about the difficulty of the procedure. If the agent is completely incompetent, $\theta = \frac{1}{2}$, then she has no prior information and thinks that both states are equally likely. A perfectly competent agent, $\theta = 1$, instead knows the correct state with certainty. The principal knows the competence level of the agent θ but he does not observe the signal s. The principal's prior about the state being 1 is 0.5.

The agent can acquire additional public information $i \in \{0, 1\}$, but the precision and the cost of this signal are proportional to the effort level. The cost of the effort level $e \in [0, 1]$ is quadratic $c(e) = ce^2$, where $c > 0$. The precision of the public signal depends on the effort level in the following way:

$$P(i = 1|\omega = 1) = P(i = 0|\omega = 0) = m(e) = 0.5(0.5e + 1),$$

$m(e)$ is the informativeness/precision of the public signal. If the effort level is 0 ($m(e) = 0.5$), then the signal is random. Since $m(e)$ is increasing in e, the probability of getting the correct public signal increases with effort. The public information acquired by the agent in the RUC example is the surveys filled out by

[4] We assume that the agent has a state-independent preference.

different doctors and the effort is the relative number of filled surveys. Observe that for $e = 1$, we have $m(e = 1) = 0.75 < 1$, meaning that even if the agent invests maximum effort in the public signal, it will never be certain.[5] This assumption is logical for our example, since no matter how many questionnaires are filled out, their informativeness is always limited (due to the structure of the questionnaires as well as the knowledge of the unknown procedure by the doctors in related fields, and their incentives to invest time and effort in being as precise as possible).

Utilities: Depending on the true state of the world and the policy choice x, players' utilities are:

$$U_p = -(\omega - x)^2;$$
$$U_A = -(1 - x).$$

These utilities show that the principal cares about "fairness" and therefore wants to implement correct prices, while the agent is biased toward the highest policy. Since the proposer's future income directly depends on the price of the medical service under consideration, she wants higher prices regardless of the real difficulty of the procedure.[6]

We characterize strategies for players in our main model where the principal chooses the effort level but cannot make an arbitrary policy decision. For this paper, we only consider pure strategies:

- The strategy of the agent is to choose private report based on the realization of her private signal $\tilde{s} : \{0, 1\} \to \{0, 1\}$.
- The principal chooses the effort level for each report from the agent $e : \{0, 1\} \to [0, 1]$.

The principal also makes a final policy decision $x(\tilde{s}, i, e(\tilde{s})) \in [0, 1]$ after observing the public signal and the private report. We consider the model where the principal cannot commit to an arbitrary decision rule for the policy choice. Rather, he has to make an interim rational policy decision based on his posterior belief about the

[5] This assumption is needed to rule out a trivial solution when the principal always requires maximum effort in the public signal and finds out the true state with certainty. The nature of the results stays the same if we vary this upper limit.

[6] Preliminary analysis shows that similar results should be achieved with a quadratic utility function for the agent as well. However, the exact calculations and closed-form solutions are easier to achieve with this utility function.

state, conditional on public information, private report, and his own beliefs. In the case of the RUC, the meetings are regulated and transparent to make sure the prices are fair. Since the goal is to match the state, the decision depends on the posterior beliefs after both private and public signals have been observed (given the effort level).

The sequence: The sequence of the game where the principal imposes formal requirements on the effort level is the following:

0) The principal commits to a mechanism requiring certain effort levels based on private reports of the agent.

1) Nature decides the state ω and private report of the agent s.

2) The agent observes her private signal s and decides what to report to the principal $\tilde{s}(s)$.

3) The principal requires the effort level from the agent, based on her private report $e(\tilde{s})$ as described in (0).

4) The agent invests $e(\tilde{s}(s))$ effort in the public signal and the public signal is realized.

5) The principal observes the public signal and a policy decision is made.

6) Utilities are realized.

It is important to notice that the agent cannot hide public information and has to invest the effort level asked. This assumption is intuitive in the RUC example, since the proposer has to collect the number of surveys required by the RUC and she cannot tamper with their results.

For comparison, we also consider the model where the agent freely chooses the effort level they want to invest in the public signal. This allows us to compare how the costly effort can be used in equilibrium by the biased agent in order to induce higher policy choices.

4.3 Preliminary Analysis: Policy Decision

Given our assumption regarding the policy choice discussed earlier, we can calculate the policy choices after each information set. Since the principal has a quadratic utility function, we already know that his interim rational policy decision is the posterior belief about the state of the world. We can use the Bayes Rule to characterize policy choices depending on each information set for different effort levels.

Observation 3 *If the principal observes public signal i and believes that the realization of the private signal is s, the policy choice follows the Bayes Rule and equals:*

$$x_{00} = P(\omega = 1|s = 0, i = 0) = \frac{(1 - m(e))(1 - \theta)}{(1 - m(e))(1 - \theta) + m(e)\theta};$$
$$x_{01} = P(\omega = 1|s = 0, i = 1) = \frac{m(e)(1 - \theta)}{m(e)(1 - \theta) + (1 - m(e))\theta};$$
$$x_{10} = P(\omega = 1|s = 1, i = 0) = \frac{(1 - (m(e))\theta}{m(e)(1 - \theta) + (1 - m(e))\theta};$$
$$x_{11} = P(\omega = 1|s = 1, i = 1) = \frac{m(e)\theta}{(1 - m(e))(1 - \theta) + m(e)\theta},$$

where $x_{j,i}$ to denotes a policy choice after observing j private and i public signals.

These policy choices are relevant in separating equilibria, where the agent chooses different effort levels depending on the realization of her private signal (and for off-path beliefs in pooling equilibria). As for the policy choice when the agent pools on the same effort level regardless of their private signal:

Observation 4 *If the agent chooses the same effort level regardless of her private signal, in equilibrium the principal disregards her private reports for this particular effort level and makes a policy decision solely based on the public signal.*

Imagine by contradiction that the agent always chooses effort level e but the principal trusts her private report $\tilde{s}$ as well. For the same public signal, the principal always makes a lower policy choice after observing a low private signal than after observing a high one: $x_{0,i} < x_{1,i}$ for $i \in \{0, 1\}$. Since the effort levels are the same, the low-type agent wants to report a high private signal and get a higher policy choice. Therefore, everyone has an incentive to report a high private signal and the principal should entirely disregard it. With a biased agent, when the agents pool on the same effort level, the private report simply becomes cheap talk.

Corollary 23 *In the pooling equilibrium, after observing low and high public signals and equilibrium effort level e, the principal's optimal policy choices are* $x_0^* = 1 - m(e)$ $\ x_1^* = m(e)$ *respectively.*

This corollary directly follows from Observation 2 using Bayes Rule.

4.4 Agent Decides Effort

We now consider the case of a biased agent who decides by herself how much effort to exert into the public signal after each realization of her private signal. This arrangement is closer to the reality of how the RUC operates. There is no official requirement for the number of surveys and the proposer decides this amount. The principal cannot commit to the decision rule and chooses his action according to his posterior belief. For the welfare comparison, we restrict our attention to equilibria that provide the highest expected payoff to the principal.

Separating Equilibrium

We consider the best separating equilibrium for the principal that satisfies the Intuitive Criterion. In separating equilibria, e_0 denotes the effort level selected by the agent with low private information and e_1 denotes the effort level selected by the agent with high private information.[7]

Observation 5 *When the agent decides effort: in a separating equilibrium,* e_0 *should maximize the low-type agent's expected utility over* $e \in [0, 1]$ *when treated as a low type.*

In a separating equilibrium, for the same effort level, the agent gets strictly more utility when considered as a high type than a low type. Therefore, regardless of the beliefs of the principal (on and off path), if e_0 does not maximize the low agent's utility when treated as a low type, she will have a profitable deviation to the maximizer. The expected utility of a low (high) type, if considered as low (high), can be written as:

$$EU_L^L = \underbrace{-\Big[P(i=0|s=0)(1-x_{00}) + P(i=1|s=0)(1-x_{01})\Big]}_{B_L^L} - ce_0^2;$$

$$EU_H^H = \underbrace{-\Big[P(i=0|s=1)(1-x_{10}) + P(i=1|s=1)(1-x_{11})\Big]}_{B_H^H} - ce_1^2.$$

The first part of the expected utilities are the benefits of effort and the second part are the costs. Observe that x_{00}, x_{01} in these equations also depends on e_0 through

[7] We refer to the private information of the agent as her type.

$m(e_0)$. The marginal cost is trivially positive $\frac{dC_L^L}{de_0} = 2ce_0 > 0$. The marginal benefits instead can be further simplified to

$$B_L^L = -\theta;$$
$$B_H^H = -(1-\theta).$$

As we can see, the marginal benefit does not depend on the level of effort in the public signal when the principal correctly assumes the agent's type. This is not surprising since from the point of view of the agent, the expected decision made by the principal (correctly believing the agent's type) equals $E(x) = \Pr(\omega = 1|s)$ and it does not vary in the precision $m(e)$ of the public signal. Moreover, in a separating equilibrium, we cannot have $e' < e_0$ such that the principal believes the agent is the high type when exerting e' level of effort. In such a case, the low-type agent would have a profitable deviation to e' being considered the high type (which results in higher policy choice for the same realizations of the public signal) and having to spend less cost on effort. These two observations imply that the low-type agent will always exert 0 effort in separating equilibrium, i.e, $e_0^n = 0$.

Since EU_H^H is also decreasing in effort, e_1 in separating equilibria should be the lowest effort level for which the principal believes that the agent is a high type; otherwise, the high agent will have a profitable deviation.

If the low type is treated as a high type by the principal, her expected utility is:

$$EU_L^H = -\left[P(i=0|s=0)(1-x_{10}) + P(i=1|s=0)(1-x_{11})\right] - ce_1^2.$$

Claim 1 *When the agent decides effort: if the low-type agent is considered to be the high type, her expected utility is decreasing in* e_1.

This claim is intuitive since the low-type agent expects the public signal to agree with her private signal and point toward the lower policy choice. Moreover, a higher effort level e_1 would both decrease the expected policy choice further and increase the cost of effort. Therefore, the low type, when considered as the high type, prefers the lowest possible effort level for e_1.

Now, in order for the low agent not to imitate the high type, we must have $EU_L^L(e_0^n) \geq EU_L^H(e_1^n)$. If this inequality is strict, a separating equilibrium will fail the Intuitive Criterion. For $e = e_1^n - \epsilon$ with $\epsilon > 0$ small enough, the principal believes that the agent is a low type, but given the most favorable strategy of the principal (considering the agent a high type), only the high agent will want to deviate to this effort level (the expected utility of both agents when considered as a high type is decreasing in effort). With this belief of the principal (considering e to come from the high agent), a high type will always want to lower her effort to e. Thus, we consider e_1^n for which the low type is indifferent $EU_L^L(e_0^n) = EU_L^H(e_1^n)$.

Proposition 24 *When the agent decides effort: a separating equilibrium only exists for high enough costs* $c > \frac{3-6\theta}{-3-4\theta+4\theta^2} \equiv \bar{c}$*. In this equilibrium, the low type exerts no effort* $e_0^n = 0$ *and the high type exerts some intermediate effort level* e_1^n*, where* e_1^n *solves* $EU_L^L(0) = EU_L^H(e_1^n)$*. The off-path beliefs are assuming the agent to be a high type i.f.f.* $e \geq e_1^n$*. This equilibrium satisfied the Intuitive Criterion.*

In the appendix, we show that there is a unique effort level e_1^* for the high type that makes the low type indifferent between reporting either private signal. We also check that the high agent does not have an incentive to deviate to any point below $e_1^n < 1$ and be perceived as a low type ($EU_H^H(e_1^n) > max_{e_0<e_1^n} EU_H^L(e_0)$).[8]

This separating equilibrium once again proves the disadvantage of letting the agent choose their own effort. When the agent can choose the level of effort they want to exert in acquiring a public signal, the agent with a low private signal exerts no effort whatsoever, while the agent who believes the true state is high exerts just high enough effort to differentiate herself from the low agent. This result agrees with empirical observation of **?** stating that in case of RUC, "the sample sizes of the survey are extraordinarily small."

When the agent chooses the effort level, the separating equilibrium only exists for a high enough cost that keeps the low agent from deviating and pretending to have a high signal just to increase the final policy choice (even exerting the maximum effort). Increasing the effort level has two disadvantages for the low type: effort is costly, the public signal is likely to decrease the policy decision even more with higher effort. When the cost is very low, $c < \bar{c}$, the first concern becomes insignificant. Moreover, since the public signal is never perfect (even with the maximum effort), sometimes the low type would rather be considered a high type, exert maximum effort and hope that the public signal disagrees with her private beliefs (as opposed to being considered a low type, exerting no effort, and getting a

[8]This condition is simplified since for $c > \frac{3-6\theta}{-3-4\theta+4\theta^2}$, the high type who is considered as low, always prefers lowest level of effort $e_0 = 0$.

low policy choice). This is the reason why the separation is not possible below the cost cutoff $\bar{c}$.

Even though both agent types have the same ex-ante preferences, what drives the separation is that after observing their private signals, their interests diverge. Mostly, divergence is due to the fact that the agents have opposite direct preferences for $m(e)$ and consequently for effort. If we fix the decision rule, the high agent's benefit is increasing in effort, while the low agent's benefit is decreasing, due to agents' expectations over whether the public signal is going to push the decision toward their preferred policy 1. Next, we characterize pooling equilibria when the agent chooses the effort level.

Pooling Equilibria

There are multiple pooling equilibria of this model supported by different off-path beliefs. For instance, one trivial pooling equilibrium is when the agent (regardless of her effort level) always exerts zero effort in the public signal. Therefore, the principal chooses a high or low policy with equal probabilities. The off-path beliefs of the principal are to assume that anyone exerting positive effort has a low private signal. Observe that this is the worst case for the principal since he receives no information (public or private) and just randomly makes the decision according to his prior belief. For comparison with the main model where the principal decides the effort levels, we will ignore these kinds of equilibria and consider only the best possible pooling equilibrium (pooling at the highest possible effort level) for the principal.

If (e, x_0^*, x_1^*) is a pooling equilibrium, the principal should believe that anyone who exerts effort lower than e has a low private signal; otherwise, both types will have a profitable deviation. Agents' utilities from pooling are:

$$EU_L^P = -\Big[P(i=0|s=0)m(e) + P(i=1|s=0)(1-m(e))\Big] - ce^2;$$
$$EU_H^P = -\Big[P(i=0|s=1)m(e) + P(i=1|s=1)(1-m(e))\Big] - ce^2.$$

The first-order condition shows that EU_L^P is strictly decreasing in effort, i.e., the low agent wants to pool at minimum effort. Similar to the separating equilibrium case, the low agent should be weakly better off than her ideal effort when perceived as a low type. In fact, since EU_L^P is decreasing and we want the best possible case for the principal (pooling at the highest effort), (1) either the low agent is indifferent between pooling and choosing her optimal effort when perceived as a low agent, or (2) pooling occurs at the highest effort level $e = 1$ (if we make sure that the high type does not want to deviate in either case). Since we have already seen that the low type's ideal effort level when considered low is 0, we can calculate the pooling effort level.

Claim 2 *When the agent decides effort: pooling has to always occur at an effort level less than or equal to e^*, with equality if there exists off-path beliefs that do not make high type want to deviate.*

$$e^* = \begin{cases} 1 \;\; if \;\; c \leq 0.375(2\theta - 1) \\ 2\sqrt{\frac{-1+2\theta}{-1+8c+2\theta}} \;\; if \;\; c > 0.375(2\theta - 1) \end{cases} .$$

Fixing the effort level, each type would rather be considered a high type than a low type. This is true because $x_{0i} < x_{1i}$ for $i \in \{0, 1\}$ for any realization of the public signal: a final policy choice is higher if the principal believes the private report was high as well. Therefore, the easiest off-path belief to sustain the pooling equilibrium is considering anyone who exerts $e \neq e^*$ to be the low type.

Proposition 25 *When the agent decides effort: in the best pooling equilibrium for the principal the agent pools at the same effort level $e^n = e^*$ described in the previous claim. The off-path beliefs are assuming the agent to be a low type. For cost low enough, $c < \bar{c}$, where the separating equilibrium does not exist, this pooling equilibrium also satisfies the Intuitive Criterion.*

As shown on the graph below, for low cost, the agent pools on the maximum effort level $e^* = 1$, and the low type is strictly better off choosing e^* over $e = 0$. In the case when pooling occurs on an intermediate effort level $e^* < 1$ (higher cost), the low-type agent is indifferent between choosing e^* or putting no effort and being considered a low type.

Proving this is actually an equilibrium strategy involves showing that the high type has no incentive to deviate and be treated as a low type. In the pooling equilibrium described above, the high type gets strictly better utility than the low type, and since e^* is chosen to make the low type indifferent, the high type does not want to deviate either.

For $c < \bar{c}$ the low-type agent is strictly better off deviating to any effort level $e \neq e^*$, if being considered a high type. This means that any off path beliefs are not equilibrium dominated by the low type and. Therefore, the off-path beliefs always assuming the agent is the low type satisfies the Intuitive Criterion.

To illustrate why the Intuitive Criterion might fail for $c > \bar{c}$, imagine an agent who deviates and exerts maximum effort level $e = 1$. Even if the low agent is considered to be the high type, she would not want to exert so much effort since (1) it is costly and (2) likely to agree with her low private signal leading to a lower policy choice. As for the high type, she actually benefits from investing in the public signal since it is likely to agree with her high private signal leading to a higher policy choice. If the cost is not too high, the high-type agent strictly wants to deviate to the maximum effort level when considered a high type, making $e = 1$ not equilibrium dominated for her. Off-path beliefs satisfying Intuitive Criterion cannot put a positive probability on the agent who exerts $e = 1$ effort being a low type. But if the principal believes $e = 1$ comes from the high agent, the high type has a profitable deviation.

To conclude, for $c > \bar{c}$, e^* is the upper bound for the best pooling equilibrium satisfying the Intuitive Criterion. However, even with this upper bound we later show that the principal strictly benefits from imposing the formal requirements on the effort level invested in the public signal.

a: Pooling at $e^n = 1$ b: Pooling at $e^n < 1$

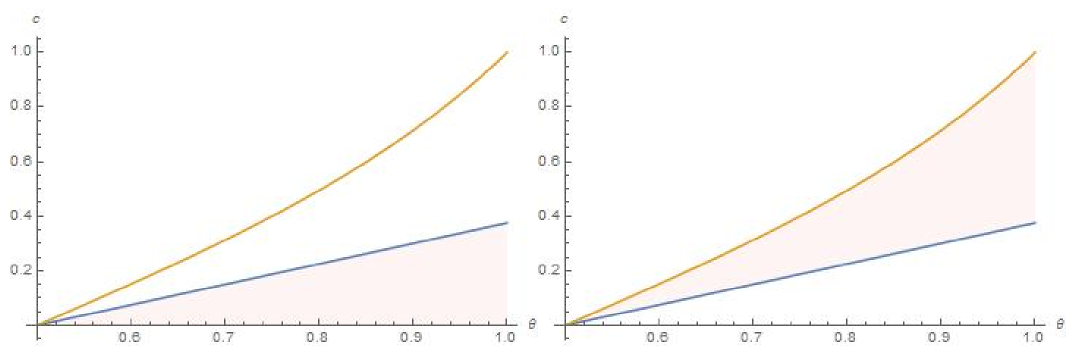


Figure 4.1: These figures illustrate the regions where the best pooling equilibrium for the principal satisfies the Intuitive Criterion and performs better than the separating equilibria. For a low cost, the agent pools at maximum effort $e = 1$ (Figure 2a). For a high cost, the agent pools at intermediate effort $e < 1$ which makes the low agent indifferent to deviation (Figure 2b). The off-path belief of the principal is always low.

4.5 No Commitment with Formal Requirements

Now we consider the case when the principal can require different effort levels from the agent, but he cannot credibly commit to an arbitrary decision rule. In the case of the RUC, it is reasonable for the committee to require a certain number of surveys depending on the soft information provided by the proposer. After providing all the soft and hard information, the committee is expected to make an "interim rational" decision. The principal's policy decision will be based on his posterior belief about the state being high, based on the reported private signal $\tilde{s}$ and public information i with required effort level $e(\tilde{s})$. Consequently, if the principal believes agent's private report, policy choices will be the same as in Observation 3: $X^* = \{x_{00}, x_{01}, x_{10}, x_{11}\}$.

The principal chooses effort levels after low and high reports from the agent. There are two possibilities:

1) Pooling equilibria: when the principal requires the same level of effort despite the agent's report $e_1 = e_2 \equiv e$.

2) Separating equilibria: when the principal requires a different level of effort based on the agent's private report $e_1 \neq e_2$.

Pooling: One Effort Level

The analysis of this case is straightforward since if the principal requires the same effort from both types, then the reports from the agents are just cheap talk. When the principal requires only one effort level, he should disregard private reports and make a decision just based on the realization of the public signal. Consequently, the principal's policy decision after observing low and high public signals will be $x_0 = 1 - m(e)$ and $x_1 = m(e)$. The objective function of the principal has the following form:

$$\max_{e \in [0,1]} \; -P(i=1)\left[P(\omega=1|i=1)(1-x_1)^2 + P(\omega=0|i=1)x_1^2\right] - \\ -P(i=0)\left[P(\omega=1|i=0)(1-x_0)^2 + P(\omega=0|i=0)x_0^2\right].$$

Simplifying this objective function and substituting the values of x_0, x_1 and $m(e)$, we get that the principal's optimal effort level for pooling.

Lemma 26 *When the principal can put a formal requirement on the effort level for public information, if he does not differentiate based on the agent's private report,* *he always requires a maximum effort level* $e^r = 1$ *and makes the policy decision solely based on the realization of the public signal.*

This is trivial since in case of pooling, the only information he takes into account is the public signal and therefore the principal wants it to be as accurate as possible.

Separating- Different Effort Levels

In this chapter, we consider only full separation- when the principal requires two different amounts of effort and agent types do not mix.[9] The objective function that the principal maximizes is:

$$\max_{e_1,e_2} -P(s=0)\Big[P(i=0|s=0)\Big(P(\omega=1|i=0,s=0)(1-x_{00})^2+P(\omega=0|i=0,s=0)x_{00}^2\Big)+$$
$$P(i=1|s=0)\Big(P(\omega=1|i=1,s=0)(1-x_{01})^2+P(\omega=0|i=1,s=0)x_{01}^2\Big)\Big]-$$
$$P(s=1)\Big[P(i=0|s=1)\Big(P(\omega=1|i=0,s=1)(1-x_{10})^2+P(\omega=0|i=0,s=1)x_{10}^2\Big)+$$
$$P(i=1|s=1)\Big(P(\omega=1|i=1,s=1)(1-x_{11})^2+P(\omega=0|i=1,s=1)x_{11}^2\Big)\Big],$$

subject to two intensive compatibility (IC) constraints for each realization of the private signal s:

$$IC_0: \quad EU_L^L(e_0) = -P(i=0|s=0)(1-x_{00}) - P(i=1|s=0)(1-x_{01}) - ce_0^2 \geq$$
$$-P(i=0|s=0)(1-x_{10}) - P(i=1|s=0)(1-x_{11}) - ce_1^2 = EU_L^H(e_1)$$
$$IC_1: \quad EU_H^H(e_1) = -P(i=0|s=1)(1-x_{10}) - P(i=1|s=1)(1-x_{11}) - ce_1^2 \geq$$
$$-P(i=0|s=1)(1-x_{00}) - P(i=1|s=1)(1-x_{01}) - ce_0^2 = EU_H^L(e_0).$$

Lemma 27 *When the principal imposes formal requirement that induce separation, for the optimal effort levels, the IC constraint of the low agent is binding and the IC constraint of the high type is slack.*

Intuitively, if IC_0 does not bind and the low agent does not exert maximum effort, then the principal would profit by increasing e_0. The low-type agent cannot exert maximum effort since this would violate her IC constraint (for both agents, it is the worst-case scenario to be considered a low type and be required to exert the maximum effort). This lemma shows that in separating equilibria, the principal

[9] Pure strategies let us concentrate on truthful revelation of the agent's type in separating equilibrium $\tilde{s} = s$.

primarily takes care of the IC constraint of the low agent, since this agent has the most incentive to hide her private information.

Corollary 28 *When the principal imposes formal requirements and induces separation, the optimal strategy requires a high-type agent to exert maximum effort, i.e.,* $e_1 = 1$.

This result follows directly from the previous lemma. Both incentive compatibility constraints cannot be binding for the given cutoffs and $e_0 \neq e_1$. Thus, IC_1 is not binding. If optimal $e_1 < 1$, increasing it with ϵ small enough does not violate any IC constraints and increases the objective function, which gives us a contradiction.

Substituting all these results in IC_0 and solving for e_0 provides the optimal solution for the case when the principal imposes different requirements based on the private report of the agent.

Proposition 29 *If the principal imposes formal requirements on the effort level for the public signal, separation is only possible for* $c > \bar{c}$*. Moreover, the optimal effort levels in this case are: maximum effort for the high type* $e_1^r = 1$ *and intermediate effort* $e_0^r = \sqrt{\frac{-3-3c+6\theta-4c\theta+4c\theta^2}{c(-3-4\theta+4\theta^2)}} < 1$ *for the low type.*

This proposition shows that, similar to the case when the agent decides the effort level, when the principal imposes formal requirements he can only induce separation for high enough cost. The effort levels have two functions for the principal: (1) more effort increases the precision of the public signal and therefore leads to better decision rules, (2) the principal uses effort requirements to satisfy the incentive compatibility constraints of the agent and extract their private information. The optimal strategy, when separation is possible, has an intuitive structure. When the agent is reporting a low private signal, she is going against her preferences and is rewarded by being required to exert less effort in the public signal. Meanwhile, the agent with a high private report has to provide maximum effort in the public signal to support her proposal. As the cost increases, formal requirements on effort become a more effective tool for inducing separation and, therefore, the low-type agent's required effort level increases as well. So far, we have characterized optimal strategies for the required effort in both separating and pooling cases. In equilibrium, the principal will evaluate his expected utility from both of these strategies and make a decision accordingly.

Equilibrium

In order to calculate the equilibrium effort levels when the agent can impose formal requirements, we compare $EU_p(e_0 = e_1 = 1)$ and $EU_p(e_0 = e_0^r, e_1^r = 1)$. Substituting the optimal values gives us the following result:

Proposition 30 *When the principal chooses the effort levels for the public signal, in equilibrium:*

- *For* $\left[\theta < \bar{\theta} \text{ and } c < \tilde{c}\right]$ *or* $\left[\theta > \bar{\theta} \text{ and } c < \bar{c}\right]$*: the principal imposes the same maximum effort level after both private reports and the policy decision is made solely based on the public signal.*
- *For* $\left[\theta \leq \bar{\theta} \text{ and } c \geq \tilde{c}\right]$ *or* $\left[\theta \geq \bar{\theta} \text{ and } c \geq \bar{c}\right]$*: the principal imposes maximum effort level after the high private report* $e_1^r = 1$ *and intermediate effort level after the low private report* $e_0^r = \sqrt{\frac{-3-3c+6\theta-4c\theta+4c\theta^2}{c(-3-4\theta+4\theta^2)}} < 1$.

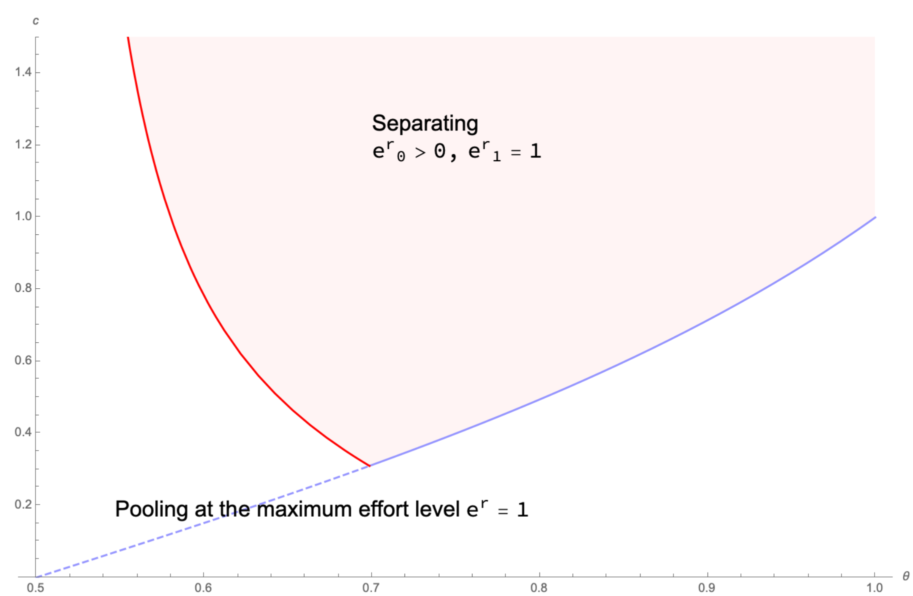


Figure 4.2: This figure shows equilibrium with formal requirements. For low enough costs, both agents exert maximum effort and the principal follows only the public signal. For high costs, the high-type agent exerts maximum effort and the low-type agent exerts intermediate effort level.

Case 1: The equilibrium for parameters on the white region of the graph has the following form:

$$e^r = 1;$$
$$x^*(e^r = 1).$$

There are two different situations in this case. (1) The cost of acquiring additional information is not high enough to satisfy the incentive compatibility of the low agent even when the high type agent is required maximum effort and the low type is required no effort ($c < \bar{c}$). Therefore, the principal disregards the reported private signal from the agent, always requiring the maximum amount of effort and making a decision based solely on this information. In the RUC example, this would mean that the cost of filling out surveys and of providing hard information is so low that it does not give the committee enough power to extract private information from the proposer. Consequently, the RUC requires the maximum number of surveys and makes a decision only based on this "hard information." (2) The cost is high enough ($c \in (\bar{c}, \tilde{c})$) to induce separation, but the separation is not worth it. When the principal induces separation, he has to incentivize the low-type agent by decreasing the required effort level for her. This means forgoing some precision in the public signal after a low private signal. For low enough θ and the cost, this "forgone" public signal is significant and it is not worth learning the private signal of a relatively incompetent agent. Therefore, even though separation is possible, the principal always requires the same maximum level of effort from the agent, and the final decision is made solely based on the public signal.

Case 2: The equilibrium for parameters on the highlighted region of the graph has the following form:

$$e_0^r < 1, \ \ e_1^r = 1.$$

This corresponds to a situation in which the cost is high enough to give leverage to the principal for extracting private information. In our example, for these parameters, the proposer truthfully reports her professional opinion. If this opinion supports high prices, the principal requires the maximum amount of public information (surveys), while the low agent is asked to exert intermediate effort. In both cases, the principal makes a final decision based on his updated posterior, after observing all the information. Note that, e_0 is increasing in cost c. This result is intuitive, since a higher cost gives more leverage to the principal and enables him to require more hard information.

Next, we do the welfare comparison between the two models to see the advantages of formal requirements for the principal.

4.6 Welfare Comparison

The graph below represents the relationship between the expected utility of the principal in equilibria of two models when the competence of the agent is $\theta = 0.65$.

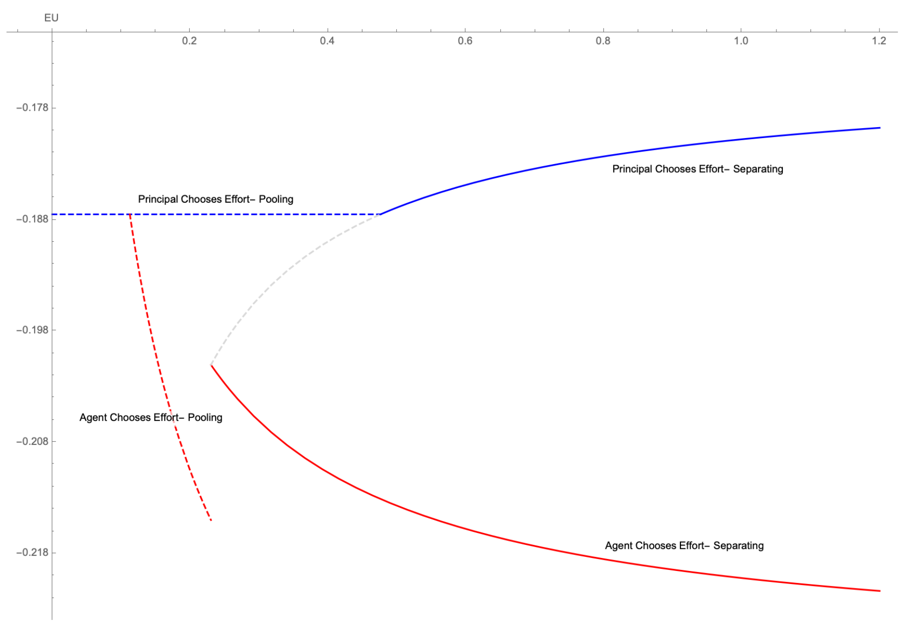


Figure 4.3: Expected Utilities of the Principal in Equilibria: $\theta = 0.65$

Proposition 31 *The principal always weakly (and sometimes strictly) benefits from imposing formal requirements on the amount of hard information acquired.*

This proposition shows that the relationship between the expected utilities shown on the graph is true for general parameters. This result has direct policy implications for the case of the RUC. As already mentioned, the current procedure puts no restriction on how many surveys the proposer should provide to support her proposal. Our analysis shows that even in the best-case scenario for the principal (our equilibrium selection favored the principal), he would achieve better results by imposing formal requirements on the number of surveys needed to support the proposal.

The proof of this proposition is rather straightforward. Separating equilibia (where the agent truthfully reveals their soft information by choosing different effort levels) in both models only exists for a cost high enough $c > \bar{c}$. When the cost is lower than this threshold, in the main model, the principal knows that he cannot impose separation (the effort is cheap, and pretending to be a high type increases the policy choice of the agent for any realization of public signal). Therefore, in equilibrium, the principal requires both types to exert maximum effort and only makes a decision based on public signal $e^r = 1$. As for the case when the agent decides the effort level, in the best equilibrium for the principal, for low enough costs, both types exert maximum effort $e^n = 1$. Since the effort is not very costly, they would rather pool on the same (maximum) effort level than be considered the low type. When the cost increases, however, the pooling effort starts to decrease $e^n < 1$ in order to make up for a higher cost and not give an incentive to the low type to deviate to no public signal. Overall, we have:

$$\text{for } c < \bar{c}, \quad e^r \geq e^n \implies EU_r \geq EU_n.$$

Since the principal always prefers a higher effort level in pooling equilibria, when both models have pooling equilibria $c < \bar{c}$, the principal is weakly better off by imposing formal requirements. Observe that when the agent controls the effort level, we select the best-case scenario for the principal. Even with this selection criterion, the principal is strictly better off with formal requirements for high enough cost $c > 0.375(2\theta - 1)$ (when $e^n < 1 = e^r$).

When the cost is high enough, $c > \bar{c}$, separation can exist in both models. However, this does not mean that when the principal controls the effort level he would always prefer to separate. In fact, he might just benefit from ignoring the private signal and always requiring the maximum level of effort. In the equilibrium of the model with formal requirement, the principal is weakly better off than if he imposed the separation with $0 < e_0^r < e_1^r = 1$. For the comparison between the two models, we can show that even with these effort levels (and separation), the principal would benefit from formal requirements on effort. When $c = \bar{c}$, both models have equilibrium with pure separation i.e., $0 = e_0 < e_1 = 1$. However, when the cost increases further, the effort levels change to:

$$e_0^r > 0, e_1^r = 1;$$
$$e_0^n = 0, e_1^n < 1.$$

Overall, we have $e_0^r > e_0^n, e_1^r > e_1^n$ and since in a separating equilibrium the principal always prefers higher efforts, he is better off with formal requirement.[10] The logic behind this result is straightforward. Both start out with the same pure separation: no effort for the low agent and maximum effort for the high agent (red solid and gray dotted lines start from the same point). When the agent decides effort, the low-type agent exerts no effort, while the high-type agent exerts just enough effort to separate themselves from the low type. Consequently, when the cost increases, the high type can more easily separate themselves even with a lower effort level and e_1^n decreases. This is the reason why the expected utility of the principal (red solid line on the graph) is decreasing in cost.

On the contrary, in the main model with formal requirements, the effort levels are used to extract soft information s but also collect the maximum possible hard information (higher effort means a more precise public signal i). When the cost increases, the effort becomes a more effective mechanism for the principal to impose separation: he requires maximum effort from the high type, but can now increase the required effort for the low type as well, without violating her incentive compatibility constraint. Therefore, the difference between the effort levels shrinks with higher cost and e_0^r slowly converges to $e_1^r = 1$. This is the reason why the expected utility of the principal (gray dotted line and blue solid line) is increasing in cost.

The reason we chose this particular value of θ is to illustrate one more interesting feature that equilibria exhibit. Sometimes, when the agents separate themselves, revealing their soft information, the principal would rather ignore it altogether and make a decision just based on the maximum required hard information (the area with the gray dotted line on the graph). This is the case when θ is small enough so the soft information is not too valuable and the cost is intermediate, so the high-type agent would use it to separate themselves from the low-type agent if they were to choose their own effort levels. In fact, when the principal imposes formal requirements, he could also induce the separation by satisfying the incentive compatibility constraints of the agents, but with the intermediate cost, it would mean letting the low type off the hook by requiring relatively lower effort (e_0^r low). This, in combination with a weak private signal (not too high θ) does not make the separation worth forgoing hard information from the low type. Therefore, the principal would rather ignore

[10]When the principal chooses the effort level, his expected utility is always weakly more than expected utility from pooling at maximum effort level $e = 1$. For this reason, the principal is weakly better off with formal requirements comparing to any pooling equilibrium of the model where the agent chooses her own effort level.

the private reports and make a decision with the maximum possible effort in hard information.

4.7 Conclusion

So far, we have shown that when the agent decides her own effort levels, bias in her preferences generates the possibility of a separating equilibrium (for a high enough cost). Even though the bias is the same for both types, observing different private signals drives a wedge between her interim preferences. The main reason for this disagreement lies in her opposite preferences for the informativeness of the public signal after observing the private signal. The low type expects the public signal to be low, leading to a low policy decision. Thus, she prefers the public signal to be as unreliable as possible. The high type, on the other hand, expects the public signal to be high, leading to a higher policy decision. Therefore, the high type prefers more reliable public signal.

Although separation is possible when the agent controls the effort level, it is still guided by biased preferences, and effort levels are not optimal for the principal. Observe that this is the case even when we select the best reasonable (satisfying Intuitive Criterion) equilibrium for the principal. In fact, if an agent's competence is low enough, they might be willing to separate themselves, but the principal would rather use maximum public signal and ignore their private reports (pool at the highest effort level).

The main purpose of this paper is to show that if the principal takes control and imposes formal requirements on the agent, he could increase the fairness of the outcome. Our analysis shows that this is in fact the case. The principal can use formal requirements on effort, not only to get a direct benefit from a better public signal, but also to correctly align the preferences of the biased agents. The optimal mechanism rewards the low private report by requiring a lower effort level for the public signal. Meanwhile, the high reports need to be accompanied by the maximum effort in the public signal. The welfare comparison shows that the principal is always weakly better off and always strictly better off when separation is possible. This result has some policy implications for our application: it shows that the current procedure is inefficient, and introducing formal requirements can significantly increase the fairness of prices.

The next step of the project involves relaxing the pure strategy constraint and characterizing the optimal mechanism for the general case. This would significantly complicate the calculations and require the application and proof of the modified version of the revelation principle, but potentially it would lead to more interesting insights into the problem.

In order to incorporate the possibility of a biased committee, a potential extension of this paper would introduce a third party (AMA) as a principal whose goal is to maximize social welfare. However, the AMA can only choose effort levels conditional on the agent's report of her private signal, while the decision about prices is still made by (potentially biased) committee members. In this case, each member of the committee would have the following utility function:

$$U_i = \alpha_i \mathbb{1}\{\tilde{d} = H\} + (1 - \alpha_i)\mathbb{1}\{\tilde{d} = \omega\},$$

where α_i measures how biased the committee member i is towards the high prices. If $\alpha_i = 0$, there is no bias and the committee member is only concerned with fairness. If $\alpha_i = 1$, the committee member only gets utility from high prices; therefore, her preference is identical to the proposer's preference in our setting. It would be interesting to observe how the expected outcome of the optimal mechanism changes with different bias levels of the committee and different voting rules.

Appendix A

CHAPTER II

Lemma 2

When the incumbent does not anticipate the external shock, the voter's strategy is the best response i.f.f. she reelects the incumbent when $\left[y = A \text{ and } u > u_A^*\right]$ or $\left[y = B \text{ and } u > u_B^*\right]$ and votes for the challenger otherwise.

$$u_A^* = \frac{1 - q - \alpha + q\alpha + q\sigma - q\alpha\sigma}{1 - q + q\sigma - \alpha\sigma}$$
$$u_B^* = \frac{\alpha - q\alpha - \alpha\sigma + q\alpha\sigma}{1 - q + q\sigma - \alpha\sigma}.$$

Proof As discussed in text, this lemma simply follows from the bias rule. After observing utility u and policy choice $y \in \{A, B\}$ posterior belief of the voter about the incumbents type is:

$$P(I = H|u, y = A) = \frac{P(u|H, y = A)P(H|y = A)}{P(u|H, y = A)P(H|y = A) + P(u|L, y = A)P(L|y = A)} =$$
$$= \frac{f(u-1)\alpha p}{f(u-1)\alpha(p + (1-p)(q + (1-q)\sigma)) + f(u)(1-\alpha)(\sigma q + (1-q))(1-p)}$$
$$P(I = H|u, y = B) = \frac{P(u|H, y = B)P(H|y = B)}{P(u|H, y = B)P(H|y = B) + P(u|L, y = B)P(L|y = B)} =$$
$$= \frac{f(u-1)(1-\alpha)p}{f(u-1)(1-\alpha)(p + (1-p)q(1-\sigma)) + f(u)\alpha(1-\sigma)(1-q)(1-p)}$$

The incumbent is reelected iff

$$P(I = H|u, y = A) > p \implies u > \frac{1 - q - \alpha + q\alpha + q\sigma - q\alpha\sigma}{1 - q + q\sigma - \alpha\sigma} \equiv u_A^*$$
$$P(I = H|u, y = B) > p \implies u > \frac{\alpha - q\alpha - \alpha\sigma + q\alpha\sigma}{1 - q + q\sigma - \alpha\sigma} \equiv u_B^*$$

Proposition 4

When the voter observes her own utility, we have two possible equilibria:

1) If $q > \tilde{q}$ truth equilibrium.

2) If $q < \tilde{q}$, the low-type politician panders after observing $s = B$ with probability $\tilde{\sigma} = \frac{(-1+q)}{(q-\alpha)} + \sqrt{\frac{(q\alpha^2 - 2q^2\alpha^2 + q^3\alpha^2 - 2q\alpha^3 + 4q^2\alpha^3 - 2q^3\alpha^3 + q\alpha^4 - 2q^2\alpha^4 + q^3\alpha^4)}{((q-\alpha)^2(q^2 + q^3 - 2q\alpha - q^2\alpha - q^3\alpha + \alpha^2 + q\alpha^2 - q^2\alpha^2 + q^3\alpha^2))}},$

where:

$$\tilde{q} = \frac{0.5(-1+2\alpha^2-\alpha^3)}{1-\alpha+\alpha^2} + 0.5\sqrt{\frac{1.+4\alpha-8\alpha^2+6\alpha^3+4\alpha^4-4\alpha^5+\alpha^6}{(1-\alpha+\alpha^2)^2}}.$$

In both types of equilibrium, the voter's strategy is to re-elect the incumbent i.f.f. ($y = A$ and $u > u^*_A$) or ($y = B$ and $u > u^*_B$).

Proof The voter's strategy in equilibrium to reelect the incumbent follows directly from Lemma 7. We have already established that H-type incumbent and L-type incumbent who sees popular private signal $s = A$, always follow their signal. Now we show that L-type incumbent who sees unpopular signal $s = B$ does not have profitable deviations in two cases considered above.

1) $q > \tilde{q}$. $\tilde{q}$ is the value for which $EU^B_B - EU^A_B$ is positive even for $\sigma = 0$ (and $EU^B_B - EU^A_B$ is increasing in σ). Therefore, for the given reelection strategy of the voter, the low-type politician who sees the unpopular signal does not have incentive to deviate and pander.

1) $q < \tilde{q}$. In this case voter's strategy is to reelect incumbent iif ($y = A$ and $u > u^*_A(\tilde{\sigma})$) or ($y = B$ and $u > u^*_B(\tilde{\sigma})$). Given this strategy of the voter, the low-type incumbent with $s = B$ signal gets utility $EU^B_B(u^*_B(\tilde{\sigma}))$ if he is truthful and $EU^A_B(u^*_A(\tilde{\sigma}))$ if he panders. Since $\tilde{\sigma}$ is chosen to equate these two expected utilities, the incumbent has no incentive to deviate.

Proposition 5

When the incumbent anticipates the external shock ϵ, we can have the following equilibrium:

1) If $q > \bar{q}$ the politician is always truthful, the voter reelects the incumbent i.f.f. ($y = A$ and $u > u_A = 1-\alpha$) or ($y = B$ and $u > u_B = \alpha$).

2) If $q < \bar{q}$:

The Incumbent:

- A low-type incumbent who receives private signal $s = B$ and anticipates a positive external shock panders with probability $\sigma^*_+(\epsilon) > 0$ when $\underline{u} < \epsilon < u_B$;
- A low-type incumbent who receives private signal $s = B$ and anticipates a negative external shock panders with probability $\sigma^*_-(\epsilon) > 0$ when $u_A - 1 < \epsilon < \bar{u} - 1$.

Otherwise the incumbent is truthful.

The Voter: After popular policy choice $y = A$:

- The voter elects the challenger for low enough utility: $V_A^* = 0$ if $u_v < u_A$;
- The voter mixes between the incumbent and the challenger with probability $V_A^* \in (0,1)$ for intermediate levels of utility $u_A < u_v < \bar{u}$;
- The voter reelects the incumbent for high enough utility: $V_A^* = 1$ if $u_v > \bar{u}$.

After unpopular policy choice $y = B$:

- The voter elects the challenger for low enough utility: $V_B^* = 0$ if $u_v < \underline{u}$;
- The voter mixes between the incumbent and the challenger with probability $V_B^* \in (0,1)$ for intermediate levels of utility $\underline{u} < u_v < u^B$;
- The voter reelects the incumbent for high enough utility: $V_B^* = 1$ if $u_v > u_B$.

Where:

$$\underline{u} = \frac{-q + 2q\alpha - \alpha^2}{-q - \alpha + 2q\alpha} \quad \bar{u} = \frac{-\alpha + \alpha^2}{-q - \alpha + 2q\alpha}$$

$$V_A^* = \begin{cases} \frac{q-\alpha}{q\alpha - q^2\alpha} & \text{for } u \in (u_A, \underline{u}) \\ \frac{0.5q^4(1-\alpha)^2 + q^3(0.5-0.5\alpha) + q^2(-0.5+0.5\alpha)\alpha}{q^3(-1+\alpha)(-0.5q-0.5\alpha+q\alpha)} & \text{for } u \in (\underline{u}, \bar{u}) \end{cases} \qquad V_B^* = \begin{cases} \frac{q-\alpha-q\alpha+q^2\alpha}{-q^2-q\alpha+2q^2\alpha} & \text{for } u \in (\underline{u}, \bar{u}) \\ \frac{q-\alpha-q\alpha+q^2\alpha}{-q\alpha+q^2\alpha} & \text{for } u \in (\bar{u}, u_B) \end{cases}$$

$$\bar{q} = \frac{0.5(-1+\alpha)}{\alpha} + 0.5\sqrt{\frac{1-2\alpha+5\alpha^2}{\alpha^2}}$$

$$\sigma_-^*(\epsilon) = \begin{cases} \frac{(p-p^2-pq+p^2q-p\alpha+p^2\alpha+pq\alpha-p^2q\alpha-p(1+\epsilon)+p^2(1+\epsilon)+pq(1+\epsilon)-p^2q(1+\epsilon))}{(-1.`p\alpha(1+\epsilon)+p^2\alpha(1+\epsilon)+pq\alpha(1+\epsilon)-p^2q\alpha(1+\epsilon))} & \text{for } \epsilon \in (u_A - 1, \underline{u} - 1) \\ \frac{(\alpha+q\alpha+\alpha^2-1.`q\alpha^2+q(1+\epsilon)-q^2(1+\epsilon)+\alpha(1+\epsilon)-3q\alpha(1+\epsilon)+2q^2\alpha(1+\epsilon))}{(-q^2(1+\epsilon)+2q^2\alpha(1+\epsilon)+\alpha^2(1+\epsilon)-2.q\alpha^2(1+\epsilon))} & \text{for } \epsilon \in (\underline{u} - 1, \bar{u} - 1) \end{cases}$$

$$\sigma_+^*(\epsilon) = \begin{cases} \frac{\alpha^3(0.25\alpha-0.25\epsilon)+q^3(0.5-\alpha)^2(-1+\epsilon)+q\alpha(\alpha(0.5`-\alpha-0.25\alpha^2)+(-0.25+0.25\alpha+0.75`\alpha^2)\epsilon)}{((0.5\alpha^2-q\alpha^2+q^2(-0.5+1\alpha))(\alpha^2(0.5-0.5\epsilon)+q(0.5-1\alpha+(-0.5+\alpha)\epsilon)))} \\ +\frac{q^2(0.25-1\alpha+0.5\alpha^2+\alpha^3+(-0.25+1.25`\alpha-1.25\alpha^2-0.5\alpha^3)\epsilon)}{((0.5\alpha^2-q\alpha^2+q^2(-0.5+1\alpha))(\alpha^2(0.5-0.5\epsilon)+q(0.5-1\alpha+(-0.5+\alpha)\epsilon)))} & \text{for } \epsilon \in (\underline{u}, \bar{u}) \\ \frac{(p\alpha-p^2\alpha-pq\alpha+p^2q\alpha-p\epsilon+p^2\epsilon+pq\epsilon-p^2q\epsilon)}{(p\alpha-p^2\alpha-pq\alpha+p^2q\alpha-p\alpha\epsilon+p^2\alpha\epsilon+pq\alpha\epsilon-p^2q\alpha\epsilon)} & \text{for } \epsilon \in (\bar{u}, u_B) \end{cases}$$

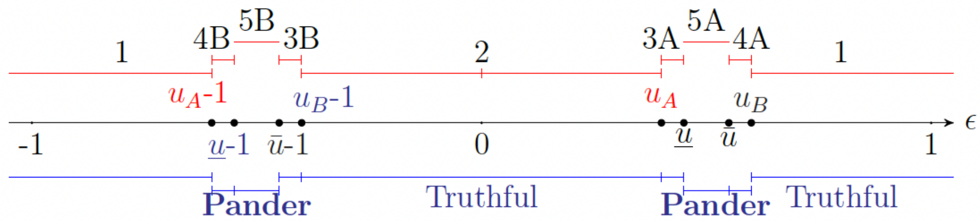


Figure A.1: *This graph shows equilibrium strategy of the low-type incumbent with unpopular signal $s = B$, depending on realization of the external shock ϵ.*

Proof Much of this proof is always done in the main text. In order to prove the given strategy profile is actually an equilibrium, we should show that neither player (the incumbent or the voter) has an incentive to deviate, given strategy of the other one.

The Incumbent

In the second period, absent reelection motives, the incumbent always follows his private signal since $q > \alpha$. We have also shown that the high-type incumbent as well as the low-type incumbent with popular signal $s = A$ are always truthful in the first-period and thus have no profitable deviation. We need to show the same for the low-type incumbent with the unpopular signal $s = B$.

Region 1 As discussed in the main text, for such extreme realizations of the external shock, the incumbent is either always reelected or always dismissed, and his policy choice has no impact on his reelection probability. Since his private signal is informative, the low-type incumbent with the unpopular signal always follows it to match the state and get a better first-period utility. In this case, the low-type incumbent has no profitable deviation.

Region 2 For small external shocks, given the equilibrium strategy of the voter, the incumbent is reelected i.f.f. he matches the state. Since his private signal is informative, the low-type incumbent is truthful and has no profitable deviation.

Region 3A For medium-low positive shocks, the incumbent is always reelected after matching the state. If the state is mismatched, he is reelected with probability V_A^* after $y = A$ and dismissed after $y = B$. In the main text, we already showed that $V'_A > V_A^*$[1] where V'_A makes the low type indifferent between being truthful

[1] V_A^* is the equilibrium reelection probability for this region.

and pandering. Since less reelection probability after $y = A$ encourages the incumbent to be truthful, he has no profitable deviation.

Region 3B For medium-low negative shocks, the incumbent is always dismissed after a mismatched state. If the state is matched, he is always reelected after the popular policy choice $y = A$ and sometimes reelected with probability V_B^* after the unpopular policy choice. We have shown in the text that $V_B^* > V_B'$[2] where V_B' makes the incumbent indifferent between being truthful and pandering. Since reelection probability after B encourages the incumbent to be truthful, he has no profitable deviation.

4A and 4B In these regions, equilibrium levels of reelection V_A and V_B are chosen to make the low-type incumbent with unpopular signal indifferent between either policy choice. Therefore, the low-type incumbent panders and has no profitable deviation.

In Region $4A$, the matched state leads to reelection of the incumbent regardless the policy choice. If, instead, the state is mismatched, then the voter always reelects the incumbent after the popular policy choice but mixes after the unpopular policy choice $y = B$ ($V_B^\epsilon < 1$). When the low-type incumbent sees signal $s = B$ if he follows it and chooses $y = B$ he gets:

$$EU_B^B = P(\omega = B|s = B)(1 + q) + P(\omega = A|s = B)(0 + V_B^\epsilon q).$$

If instead he panders and chooses $y = A$ he gets:

$$EU_B^A = P(\omega = A|s = B)(1 + q) + P(\omega = B|s = B)(0 + q).$$

Since the low-type incumbent panders in this region, the equilibrium reelection probability V_B^* makes him indifferent between following the unpopular policy and pandering, meaning:

$$V_B^* = \frac{P(\omega = A|s = B)(1 + q) - P(\omega = B|s = B)}{P(\omega = A|s = B)q}.$$

We have already shown that this reelection probability makes the low-type incumbent truthful in Region $3B$.

In Region $4B$, the voter never reelects the incumbent regardless of the policy choice after a mismatched state. If the state is matched then he voter sometimes

[2] V_B^* is the equilibrium reelection probability for this region.

reelects the incumbent after the popular policy choice $y = A$ ($V_A^{1+\epsilon} < 1$) but always elects the challenger after the unpopular policy choice $y = B$. When the low-type incumbent sees the unpopular signal $s = B$, if he follows it and chooses $y = B$ he gets:

$$EU_B^B = P(\omega = B|s = B)(1 + 0) + P(\omega = A|s = B)(0 + 0).$$

If instead he panders and chooses $y = A$ he gets:

$$EU_B^A = P(\omega = A|s = B)(1 + V_A^{1+\epsilon} q) + P(\omega = B|s = B)(0 + 0).$$

Since the low-type incumbent with the unpopular signal panders in this region, the equilibrium reelection probability is derived by equating the two expressions above:

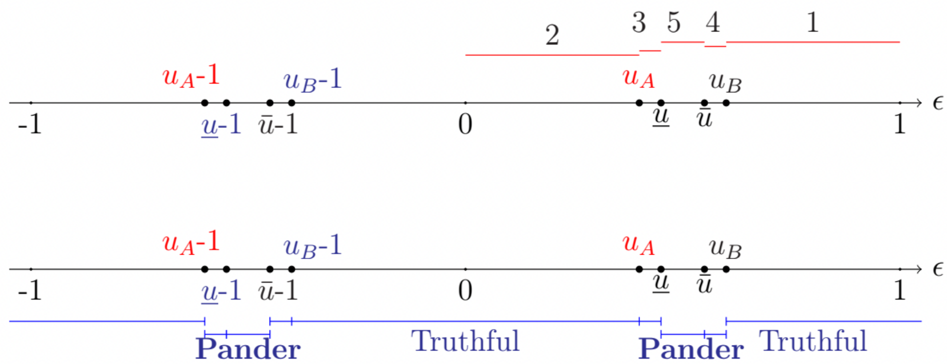
$$V_A^* = \frac{P(\omega = B|s = B) - P(\omega = A|s = B)}{P(\omega = A|s = B)q}.$$

We have already shown that this reelection probability makes the low-type incumbent truthful in Region $3A$.

5A and 5B Similar to the previous case, reelection mixing probabilities make the low-type incumbent with unpopular signal indifferent between either policy choice and he has no profitable deviation. Unlike the previous case, both reelection probabilities play a role in satisfying the indifferent conditions for the low-type incumbent with the unpopular signal in both of these regions. Consequently, the equilibrium values of the reelection probabilities (V_A^*, V_B^*) for these regions are derived by solving the system of equations for double indifference as shown in the main text.

The Voter

As already discussed in the main text, voter who observes utility beyond $(0, 1)$ range has no profitable deviation since she learns about matched/mismatched state (given challenger and incumbent are from the same pool, this is enough information for them to elect a second-period representative). The rest of the proof falls directly from the observations in Section 5.3 and uses Bayes Rule for updated posterior beliefs of the voter.



As discussed in the main text, the voter knows the low-type incumbent was truthful when she observes the utility level from Regions 1 and 2. Since $u_A = 1 - \alpha$ and $u_B = \alpha$, it follows directly from the truthful equilibrium of the baseline model that she has no profitable deviation.

For utility levels from Region 3, the voter knows that one of the two cases is possible:

- The state was matched and $\epsilon = u_v - 1$. In which case, $u_A - 1 < \epsilon < \underline{u} - 1$ and from the previous part, the incumbent is pandering with probability $\sigma^*_{\underline{ }}$.
- The state was mismatched and $\epsilon = u_v$. In which case, $u_A < \epsilon < \underline{u}$, and the L-type incumbent is truthful.

Based on this information the voter's posterior beliefs after each policy choice are:

$$P(H|u, y = A) = \frac{f(u-1)p\alpha}{f(u-1)\alpha p + f(u-1)\alpha(1-p)(q + (1-q)\sigma^*_{\underline{ }}) + (1-\alpha)f(u)(1-p)(1-q)};$$

$$P(H|u, y = B) = \frac{f(u-1)p(1-\alpha)}{f(u-1)(1-\alpha)p + f(u-1)(1-\alpha)(1-p)q(1-\sigma^*_{\underline{ }}) + \alpha f(u)(1-q)(1-p)}.$$

The equilibrium level of pandering $\sigma^*_{\underline{ }}(\epsilon = u_v - 1)$ in this region is derived to make the voter indifferent between choosing the incumbent or the challenger after observing the popular policy choice ($P(H|u, y = A) = p$). We also make sure that, for this equilibrium level of pandering, the voter wants to elect the challenger after observing the unpopular policy choice $y = B$ and the utility from this region ($P(H|u, y = B) < p$). This derives the threshold $\underline{u}$ in equilibrium.

In Region 4, the voter knows that one of the two cases is possible:

- The state was matched and $\epsilon = u_v - 1$. In which case, $\bar{u} - 1 < \epsilon < u_B - 1$ and from the previous part, the incumbent is truthful.

- The state was mismatched and $\epsilon = u_v$. In which case, $\bar{u} < \epsilon < u_B$, and the L-type incumbent is pandering with probability σ_+^*.

Based on these beliefs, the voter's posterior beliefs after each policy choice are:

$$P(H|u, y = A) = \frac{f(u-1)p\alpha}{f(u-1)\alpha p + f(u-1)\alpha(1-p)q + (1-\alpha)f(u)(\sigma_+^* q + (1-q))(1-p)};$$
$$P(H|u, y = B) = \frac{f(u-1)p(1-\alpha)}{f(u-1)(1-\alpha)p + f(u-1)(1-\alpha)(1-p)q + \alpha f(u)(1-q)(1-\sigma_+^*)(1-p)}.$$

In equilibrium, the voter is indifferent between voting for the incumbent or the challenger after observing the unpopular policy choice $y = B$ and the utility level from this region. Therefore, the equilibrium level of pandering $\sigma_+^*(\epsilon = u_v)$ is chosen to satisfy this indifference condition $P(H|u, y = B) = p$. The voter has no profitable deviation if she does not want to vote for the challenger after observing the popular policy choice $y = A$ for this equilibrium level of pandering $\sigma_+^*(\epsilon = u_v)$. Therefore we must have $P(H|u, y = A) > p$. This condition is always satisfied for $u > \bar{u}$.

For Region 5 we have already shown in the main text how the indifference conditions for the voter after both policy choices define the equilibrium levels of pandering (σ_-^*, σ_+^*) for the corresponding values of the external shock. This completes our proof.

Appendix B

CHAPTER III

B.1 Preliminary Analysis

In this appendix we conduct a general preliminary analysis of the model; the proof of main text Lemma 9 characterizing a voter best response is contained herein.

To more easily accommodate ex-ante agnosticism as to whether a low-ability incumbent distorts his policymaking toward the popular policy A or the unpopular policy B in equilibrium, we rewrite a low-ability incumbent's strategy as $\eta = (\eta^A, \eta^B)$, where η^x for $x \in \{A, B\}$ denotes the probability that the incumbent chooses policy $y = x$ after receiving signal $s = \neg x$. Hence, using our main text notation $\eta^A = \theta_B$ is the probability of "pandering" and $\eta^B = 1 - \theta_A$ is the probability of "fake leadership." We also use $\theta = (\theta^A, \theta^B)$ to denote the entire vector of a voter strategy, where $\theta^x = (v^x_\emptyset, \rho^x, v^x_x, v^x_{\neg x})$ for $x \in \{A, B\}$.

The Incumbent's Problem To formally characterize a low-ability incumbent's best responses, we first introduce notation to describe the electoral consequences of choosing each policy $x \in \{A, B\}$ given a voter strategy θ. Let

$$v^x_{\mathcal{I}}(\theta^x) = (1 - \rho^x)v^x_\emptyset + \rho^x \left(P(\omega = x|\mathcal{I})v^x_x + P(\omega \neq x|\mathcal{I})v^x_{\neg x}\right)$$

denote a low-ability incumbent's expected probability of reelection after choosing $x \in \{A, B\}$ when he has information $\mathcal{I}$ about the state and the voter uses strategy θ^x in response to first-period policy x. Applying the notation in the main text we have $EU^x_{\mathcal{I}} = P(\omega = x|\mathcal{I}) + \delta q \cdot v^x(\mathcal{I}; \theta^x)$. Next, let $\Delta^x_{\mathcal{I}}(\theta) = v^x_{\mathcal{I}}(\theta^x) - v^{\neg x}_{\mathcal{I}}(\theta^{\neg x})$ denote a low-ability incumbent's *net gain* in the probability of reelection from choosing x vs. $\neg x$ when he has information $\mathcal{I}$ and the voter uses strategy $\theta = (\theta^x, \theta^{\neg x})$. Finally, let

$$\bar{\Delta}^x_{\mathcal{I}} = \frac{\Pr(\omega = \neg x|\mathcal{I}) - \Pr(\omega = x|\mathcal{I})}{\delta q},$$

and observe that $\bar{\Delta}^x_{s=\neg x} > 0\ \forall x \in \{A, B\}$ since $q > \pi$. A low-ability incumbent's best-response is then as follows.

Lemma 32 *A low-ability incumbent's strategy $\eta = (\eta^A, \eta^B)$ is a best response to θ i.f.f.*

$$\Delta^x_{s=\neg x}(\theta) > (<) \bar{\Delta}^x_{s=\neg x} \rightarrow \eta^x = 1(0)\ \forall x \in \{A, B\}$$

.

Proof: Straightforward and omitted. QED.

The Voter's Problem When the voter is initially called to play, she has observed the incumbent's first-period policy choice x, and must choose her likelihood of paying attention ρ^x and of retaining the $v^x_\emptyset$ incumbent should she choose not to pay attention. Should she choose to pay attention, she then anticipates learning the state ω and deciding on the likelihood of retaining the incumbent v^x_ω conditional on this additional information.

We first discuss the voter's belief formation. Although some sequences of play may be off the path of play given a low-ability incumbent's strategy (for example, failure of a policy x when a low-ability incumbent is believed to always choose $\neg x$) it is easily verified that sequentially consistent beliefs about the incumbent's ability $v^x_\emptyset$ and the state $P(\omega = x|y = x)$ prior to the attentional decision ρ^x, as well as sequentially consistent beliefs μ^x_ω for $\omega \in \{A, B\}$ about the incumbent's ability after paying attention, are all unique and straightforwardly characterized by Bayes' Rule (as described in the main text). We begin with two useful algebraic equalities about these beliefs.

Lemma 33 $\Pr(\omega = x|y = x) \cdot \mu^x_x = \mu^x$

Proof: $\Pr(\omega = x|y = x) \cdot \mu^x_x$

$$\begin{aligned}
&= \frac{\Pr(y = x, \omega = x)}{\Pr(y = x)} \cdot \Pr(\lambda_I = H|y = x, \omega = x) = \frac{\Pr(\lambda_I = H, y = x, \omega = x)}{\Pr(y = x)} \\
&= \frac{\Pr(y = x|\lambda_I = H, \omega = x)\Pr(\omega = x) \cdot \Pr(\lambda_I = H)}{\Pr(y = x)} \\
&= \frac{(\Pr(y = x|\lambda_I = H, \omega = x)\Pr(\omega = x) + \Pr(y = x|\lambda_I = H, \omega \neq x)\Pr(\omega \neq x)) \cdot \Pr(\lambda_I = H)}{\Pr(y = x)} \\
&= \frac{\Pr(y = x|\lambda_I = H) \cdot \Pr(\lambda_I = H)}{\Pr(y = x)} = \mu^x,
\end{aligned}$$

where the second-to-last equality follows from $\Pr(y = x|\lambda_I = H, \omega \neq x) = 0$. QED.

Lemma 34 $\mu^x = \Pr(\omega = x|y = x)\mu_x^x + \Pr(\omega = \neg x|y = x)\mu_{\neg x}^x$

Proof:

$$\begin{aligned}
\mu^x &= \frac{\Pr(\lambda_I = H, y = x)}{\Pr(y = x)} = \frac{\Pr(\lambda_I = H, y = x, \omega = x) + \Pr(\lambda_I = H, y = x, \omega \neq x)}{\Pr(y = x)} \\
&= \frac{\Pr(\omega = x, y = x)\Pr(\lambda_I = H|\omega = x, y = x)}{\Pr(y = x)} + \frac{\Pr(\omega \neq x, y = x)\Pr(\lambda_I = H|\omega \neq x, y = x)}{\Pr(y = x)} \\
&= \Pr(\omega = x|y = x)\,\mu_x^x + \Pr(\omega \neq x|y = x)\,\mu_{\neg x}^x \qquad \text{QED}
\end{aligned}$$

With these beliefs in hand, it is easily verified that after observing first period policy $y = x$, the voter's expected utility from her anticipated strategy $\theta^x = (\nu_\emptyset^x, \rho^x, \nu_x^x, \nu_{\neg x}^x)$ following policy x is equal to:

$$V(\theta^x|\eta) = \delta q + \delta(1-q)\begin{pmatrix} (1-\rho^x)\left(\nu_\emptyset^x \mu^x + \left(1-\nu_\emptyset^x\right)\gamma\right) \\ +\rho^x \begin{pmatrix} \Pr(\omega \neq x|y = x)\left(\nu_{\neg x}^x \mu_{\neg x}^x + \left(1-\nu_{\neg x}^x\right)\gamma\right) \\ +\Pr(\omega = x|y = x)\left(\nu_x^x \mu_x^x + \left(1-\nu_x^x\right)\gamma\right) \end{pmatrix} \end{pmatrix} - \rho^x c,$$

where the unique sequentially-consistent values of $\left(\mu^x, \mu_x^x, \mu_{\neg x}^x, \Pr(\omega = x|y = x)\right)$ depend on a low-ability incumbent's strategy η.

It is next immediate that the voter's retention probabilities ν_s^x after $s \in \{\emptyset, x, \neg x\}$ (where $s = \emptyset$ denotes the decision to pay no attention and learn nothing about the state) will be sequentially rational if and only if $\mu_s^x > (<)\,\gamma \rightarrow \nu_s^x = 1(0)$. To examine the voter's attention decision ρ^x, recall from the main text that the values of negative and positive attention $\left(\phi_-^x, \phi_+^x\right)$ following policy x are defined to be:

$$\phi_-^x = \delta(1-q)\cdot\Pr(\omega \neq x|y = x)\left(\gamma - \mu_{\neg x}^x\right);$$

$$\phi_+^x = \delta(1-q)\cdot\Pr(\omega = x|y = x)\left(\mu_x^x - \gamma\right).$$

It is readily apparent that ϕ_-^x is strictly increasing in γ (ceteris paribus) while ϕ_+^x is strictly decreasing in γ (ceteris paribus). The following lemma helps connect these values to the voter's expected utility.

Lemma 35 $\mu^x - \gamma = \frac{1}{\delta(1-q)}\left(\phi_+^x - \phi_-^x\right)$

Proof:

$$\begin{aligned}\mu^x - \gamma &= \left(\Pr(\omega = x|y = x)\,\mu_x^x + \Pr(\omega \neq x|y = x)\,\mu_{\neg x}^x\right) - \gamma \\ &= \Pr(\omega = x|y = x)\left(\mu_x^x - \gamma\right) - \Pr(\omega \neq x|y = x)\left(\gamma - \mu_{\neg x}^x\right) \\ &= \frac{\phi_+^x - \phi_-^x}{\delta(1-q)}. \qquad \text{QED.}\end{aligned}$$

Finally, the following facilitates comparisons between the values of information across policies that will be useful later in the analysis.

Lemma 36 $\phi_+^{\neg x} > (=)\, \phi_-^x \iff$

$$\frac{\mu - \Pr(y = \neg x|\omega = \neg x)\,\gamma}{\Pr(y = \neg x)} > (=)\, \frac{\Pr(y = x|\omega = \neg x)\,\gamma}{\Pr(y = x)}$$

Proof: Observe from the definitions that $\phi_+^{\neg x} > (=) \phi_-^x \iff$

$$\Pr(\omega = \neg x|y = \neg x)\left(\mu_{\neg x}^{\neg x} - \gamma\right) > (=)\, \Pr(\omega = \neg x|y = x)\,\gamma$$

We first transform the lhs; we have that $\Pr(\omega = \neg x|y = \neg x)\left(\mu_{\neg x}^{\neg x} - \gamma\right) =$

$$\begin{aligned}&\mu^{\neg x} - \Pr(\omega = \neg x|y = \neg x)\cdot\gamma \text{ (using Lemma 33)} \\ =\ &\frac{\Pr(\omega = \neg x)}{\Pr(y = \neg x)}\left(\mu - \Pr(y = \neg x|\omega = \neg x)\,\gamma\right) \text{ (using } \Pr(y = \neg x|\lambda_I = H) = \Pr(\omega = \neg x)\text{)}\end{aligned}$$

We next transform the rhs; we have that $\Pr(\omega = \neg x|y = x)\,\gamma = \frac{\Pr(\omega=\neg x)}{\Pr(y=x)} \Pr(y = x|\omega = \neg x)\,\gamma$. Substituting in and rearranging then yields the desired condition. QED

With Lemmas 33-36 in hand, imposing sequential rationality on each v_s^x and rearranging yields that the voter's expected utility $V(\rho^x|\eta)$ conditional on ρ^x is equal to:

$$V(\rho^x|\eta) = \delta q + \delta(1-q)\max\{\mu^x, \gamma\} + \rho^x\left(\max\left\{\min\left\{\phi_-^x, \phi_+^x\right\}, 0\right\} - c\right).$$

This immediately yields main text Lemma 9 characterizing necessary and sufficient conditions for a voter strategy θ^x following x to be a best-response (where "best response" is used colloquially to mean "sequentially rational given the unique sequentially-consistent beliefs implied by the incumbent's strategy"). We restate Lemma 9 here, letting $\bar{\Theta}^x(\eta)$ denote the set of best responses following x when a low-ability incumbent uses strategy η.

Lemma 1 (restated) $\hat{\theta}^x$ is a best-response following $x \iff$

$\hat{v}^x_{\neg x} = 0, \mu^x_s > (<) \, \gamma \rightarrow \hat{v}^x_s = 1(0) \;\; \forall s \in \{\emptyset, x\},$ and $c < (>) \, \phi^x = \min\{\phi^x_-, \phi^x_+\} \rightarrow \hat{\rho}^x = 1 \, (0).$

Properties of Equilibrium We conclude this section by proving some basic properties of equilibrium and providing an intermediate characterization. The first property states that equilibrium may involve pandering *or* fake leadership, but not both.

Lemma 37 *In equilibrium,* $\eta^x > 0$ *for at most one* x*.*

Proof: First observe that $\eta^x > 0$ (the incumbent panders toward x) $\rightarrow EU^x_{s=\neg x} \geq EU^x_{s=x}$ (the incumbent benefits from choosing x even after signal $\neg x$) $\rightarrow v^x_{s=\neg x}(\theta) > v^{\neg x}_{s=\neg x}(\theta)$ (choosing x is electorally advantageous after signal $s = \neg x$) since $P(\omega = \neg x | s = \neg x) > P(\omega = x | s = \neg x) > 0$ ($\neg x$ is strictly more likely to be correct than x following signal $s = \neg x$ from $q > \pi$). Next observe that $v^x_{s=\neg x}(\theta) > v^{\neg x}_{s=\neg x}(\theta) \rightarrow v^x_{s=x}(\theta) > v^{\neg x}_{s=x}(\theta)$ (if x is electorally advantageous after signal $s = \neg x$ then it remains electorally advantageous after signal $s = x$), since $(v^x_{s=x}(\theta) - v^{\neg x}_{s=x}(\theta)) - (v^x_{s=\neg x}(\theta) - v^{\neg x}_{s=\neg x}(\theta)) =$

$$\rho^x \cdot (P(\omega = x | s = x) - P(\omega = x | s = \neg x)) \cdot \left(v^x_x - v^x_{\neg x}\right)$$
$$+\rho^{\neg x} \cdot (P(\omega = \neg x | s = \neg x) - P(\omega = \neg x | s = x)) \cdot \left(v^{\neg x}_{\neg x} - v^{\neg x}_x\right),$$

which is ≥ 0 since $v^x_x \geq v^x_{\neg x}$ in any best response (the voter is more likely to reelect after observing a match than after observing a mismatch) and $P(\omega = x | s = x) > P(\omega = x | s = \neg x))$ (x is more likely to be correct following signal $s = x$ than signal $s = \neg x$, by $q > \frac{1}{2}$). Finally, the preceding immediately yields $EU^x_{s=x} > EU^x_{s\neg=x} \rightarrow \eta^{\neg x} = 0$ since $P(\omega = x | s = x) > P(\omega = \neg x | s = x) > 0$ (x is strictly more likely to be correct than $\neg x$ following signal $s = x$, again by $q > \pi$). QED.

The second property states that any equilibrium involving a distortion must be mixed.

Lemma 38 *If* $\eta^x > 0$ *then* $\eta^x < 1$.

Proof: Suppose $\eta^x = 1$ (so $\eta^{\neg x} = 0$). Then $\mu^{\neg x} = 1$ and $\phi^{\neg x}_- = 0$, so a voter's best-response requires $v^{\neg x}_\emptyset = 1$ and $\rho^{\neg x} = 0$, implying $v^{\neg x}_{\mathcal{I}}(\theta) = 1 \geq v^x_{\mathcal{I}}(\theta)$. Since also $P(\omega = \neg x|s = \neg x) > P(\omega = x|s = \neg x)$ we must have $EU^{y=\neg x}_{s=\neg x} > EU^{y=x}_{s=\neg x}$, and any $\eta^x > 0$ cannot be an incumbent best-response. QED.

Collecting the preceding results, yields an intermediate characterization of equilibrium as a corollary.

Corollary 39 *Profile $(\hat{\eta}, \hat{\theta})$ is a sequential equilibrium i.f.f. it satisfies Lemma 9 and either*

- $\hat{\eta}^x = 0$ *and* $\Delta^x_{s=\neg x}(\theta) \leq \bar{\Delta}^x_{s=\neg x}$ $\forall x \in \{A, B\}$ *(the incumbent is truthful),*
- $\exists z$ *s.t.* $\hat{\eta}^z \in (0, 1)$, $\hat{\eta}^{\neg z} = 0$, *and* $\Delta^z_{s=\neg z}(\theta) = \bar{\Delta}^z_{s=\neg z}$ *(the incumbent distorts toward z).*

B.2 Equilibrium Characterization

Herein we continue the equilibrium analysis including proofs of Propositions 10 and 11. We first examine properties of the values of attention when the incumbent is truthful.

Lemma 40 *Let $\bar{\phi}^x_s$ denote the values of attention when a low-ability incumbent is truthful and $\bar{\phi}^x = \min\{\bar{\phi}^x_-, \bar{\phi}^x_+\}$. These values satisfy the following properties:*

- $\bar{\phi}^A_+ > \bar{\phi}^B_+$ *and* $\bar{\phi}^A_- < \bar{\phi}^B_-$;
- $\bar{\phi}^B > \bar{\phi}^A \rightarrow \gamma < \bar{\mu}^A$;
- $\bar{\phi}^A > \bar{\phi}^B \rightarrow \gamma > \mu$.

Proof: From the definitions, $\phi^B_- > \phi^A_- \iff \Pr(\omega = A|y = B) > \Pr(\omega = B|y = A) \iff$

$$\left(\frac{\Pr(y = A|\omega = A)}{\Pr(y = A|\omega = B)}\right)\left(\frac{\Pr(\omega = A)}{1 - \Pr(\omega = A)}\right) > \left(\frac{\Pr(y = B|\omega = B)}{\Pr(y = B|\omega = A)}\right)\left(\frac{1 - \Pr(\omega = A)}{\Pr(\omega = A)}\right).$$

When a low-ability incumbent is truthful, $\frac{\Pr(y=A|\omega=A)}{\Pr(y=A|\omega=B)} = \frac{\mu+(1-\mu)q}{(1-\mu)(1-q)} = \frac{\Pr(y=B|\omega=B)}{\Pr(y=B|\omega=A)}$, so the condition reduces to $\Pr(\omega = A) = \pi > \frac{1}{2}$. Next, from the definitions $\phi^A_+ > (<)(=)\phi^B_+ \iff \Pr(\omega = A|y = A) > (<)(=)\Pr(\omega = B|y = B)$ when a low-ability incumbent is truthful (using that $\bar{\mu}^A_A = \bar{\mu}^B_B$) which in turn holds $\iff$ $\Pr(\omega = A|y = B) > \Pr(\omega = B|y = A)$, which is already shown.

The statement that $\bar{\phi}^B > \bar{\phi}^A \rightarrow \gamma < \bar{\mu}^A$ follows trivially from the first property.

The final property is equivalent to $\gamma \leq \mu \rightarrow \bar{\phi}^B \geq \bar{\phi}^A$. To show this we argue that $\bar{\phi}^B_+(\mu) > \bar{\phi}^A_-(\mu)$. From this it is easy to verify the desired property using that (i) $\mu \in (\bar{\mu}^B, \bar{\mu}^A)$, (ii) $\bar{\phi}^B_- > \bar{\phi}^A_-$, (iii) $\phi^x_-(\gamma)$ decreasing in γ, and (iv) $\phi^x_+(\gamma)$ increasing in γ. First observe from Lemma 36 that for any values of (σ, γ) we have $\phi^B_+ > \phi^A_-$ i.f.f.

$$\Pr(y = A) \cdot \left(\gamma - \frac{\gamma - \mu}{\Pr(y = A|\omega = B)}\right) > \Pr(y = B) \cdot \gamma$$

. Next observe that when $\gamma = \mu$ the condition reduces to $\Pr(y = A) > \Pr(y = B)$, which always holds when a low-ability incumbent is truthful. QED.

We next examine how a low-ability incumbent's potential distortions η affects these values of attention. Our next two lemmas are used to this end.

Lemma 41 $\Pr(\omega \neq x|y = x)$ *is strictly increasing in* η^x *(when* $\eta^{\neg x}$*=0) and strictly decreasing in* $\eta^{\neg x}$ *(when* $\eta^x = 0$).

Proof:

$$\begin{aligned}\Pr(\omega \neq x|y = x) &= \frac{\Pr(y = x|\omega \neq x) \cdot (1 - \pi^x)}{\Pr(y = x|\omega = x) \cdot \pi^x + \Pr(y = x|\omega \neq x) \cdot (1 - \pi^x)} \\ &= \frac{1}{\frac{\Pr(y=x|\omega=x)}{\Pr(y=x|\omega\neq x)} \cdot \frac{\pi^x}{1-\pi^x} + 1}.\end{aligned}$$

So η^x ($\eta^{\neg x}$) affect the desired quantity solely through $\frac{\Pr(y=x|\omega=x)}{\Pr(y=x|\omega\neq x)}$, where:

$$\frac{\Pr(y = x|\omega = x)}{\Pr(y = x|\omega \neq x)} = \frac{\mu + (1 - \mu) \cdot (q(1 - \eta^{\neg x}) + (1 - q)\eta^x)}{(1 - \mu) \cdot ((1 - q)(1 - \eta^{\neg x}) + q\eta^x)}.$$

To perform comparative statics η^x, assume $\eta^{\neg x} = 0$ so

$$\begin{aligned}\frac{\Pr(y = x|\omega = x)}{\Pr(y = x|\omega \neq x)} &= \frac{\mu + (1 - \mu) \cdot (q + (1 - q)\eta^x)}{(1 - \mu) \cdot ((1 - q) + q\eta^x)} \\ &= \frac{\mu + (1 - \mu) \cdot (1 - q(1 - \eta^x) + (2q - 1)(1 - \eta^x))}{(1 - \mu) \cdot (1 - q(1 - \eta^x))} \\ &= 1 + \left(\frac{\mu}{1 - \mu}\right)\left(\frac{1}{1 - q(1 - \eta^x)}\right) + \frac{(2q - 1)(1 - \eta^x)}{1 - q(1 - \eta^x)}\end{aligned}$$

which is straightforwardly decreasing in η^x when $q \geq \frac{1}{2}$.

To perform comparative statics in $\eta^{\neg x}$, assume that $\eta^x = 0$ so

$$\frac{\Pr(y = x|\omega = x)}{\Pr(y = x|\omega \neq x)} = \frac{\mu + (1 - \mu) q (1 - \eta^{\neg x})}{(1 - \mu) \cdot (1 - q)(1 - \eta^{\neg x})} = \frac{\frac{\mu}{1-\eta^{\neg x}} + (1 - \mu) q}{(1 - \mu)(1 - q)}$$

which is clearly strictly increasing in $\eta^{\neg x}$. QED.

Lemma 42 $\Pr(\omega = x|y = x)(\mu_x^x - \gamma)$ *is strictly decreasing in* η^x *(when* $\eta^{\neg x} = 0$*) and strictly increasing in* $\eta^{\neg x}$ *(when* $\eta^x = 0$*).*

Proof: First observe that $\Pr(\omega = x|y = x)$ is strictly decreasing (increasing) in η^x ($\eta^{\neg x}$) by Lemma 41. Next

$$\mu_x^x = \frac{\mu}{\mu + (1-\mu)\left(q\left(1-\eta^{\neg x}\right) + (1-q)\,\eta^x\right)},$$

which is also straightforwardly strictly decreasing (increasing) in η^x ($\eta^{\neg x}$). QED.

The preceding lemmas immediately yield comparative statics effects of $\eta^x \geq 0$ (when $\eta^{\neg x} = 0$) on the four relevant values of information ($\phi_-^x, \phi_+^x, \phi_-^{\neg x}, \phi_+^{\neg x}$) as a corollary.

Corollary 43 *Suppose that* $\eta^{\neg x} = 0$*. Then* $\phi_-^x(\eta^x)$ *and* $\phi_+^{\neg x}(\eta^x)$ *are strictly increasing in* η^x*, while* $\phi_+^x(\eta^x)$ *and* $\phi_-^{\neg x}(\eta^x)$ *are strictly decreasing in* η^x*.*

We now use the preceding results to examine how an anticipated distortion $\eta^z > 0$ toward some policy z (with $\eta^{\neg z} = 0$) affects the *electoral incentives* of a low-ability incumbent when the voter best-responds. This analysis yields a key lemma which implies that the model is well behaved. The lemma states that (despite the greater complexity of the RA model), a greater distortion toward some policy z still makes that policy relatively less electorally appealing once the voter best responds (as in the CHS model). To state the lemma formally, let

$$\mathbf{\Delta}_{\mathcal{I}}^z\left(\eta^z\right) = \left\{\Delta :\ \exists\, \theta \text{ satisfying } \theta^x \in \bar{\Theta}^x\left(\eta^x\right)\ \forall x \in \{A, B\} \text{ and } \Delta = \Delta_{\mathcal{I}}^z\left(\theta\right)\right\}$$

denote the **set of reelection probability differences** from choosing policy z vs. policy $\neg z$ for a low-ability incumbent with information $\mathcal{I}$ that can be generated by a voter best response to $\eta^z \in [0, 1]$ (with $\eta^{\neg z} = 0$).

Lemma 44 $\mathbf{\Delta}_{\mathcal{I}}^z\left(\eta^z\right)$ *is an upper-hemi continuous, compact, convex-valued,* ***decreasing*** *correspondence that is constant and singleton everywhere except at (at most) four points.*

Proof: Starting with the voter's objective functions $V(\theta^x|\eta)$ and the best responses stated in main text Lemma 9 and Appendix Lemma 32, it is straightforward to verify all properties of the correspondence except that it is decreasing using standard arguments.

To argue that $\mathbf{\Delta}^z_{\mathcal{I}}(\eta^z)$ is decreasing, first observe that:

$\mathbf{\Delta}^z_{\mathcal{I}}(\eta^z) = \mathbf{V}^z_{\mathcal{I}}(\eta^z) - \mathbf{V}^{\neg z}_{\mathcal{I}}(\eta^z)$, where $\mathbf{V}^x_{\mathcal{I}}(\eta^z) = \{v : \exists \theta^x \in \bar{\Theta}(\eta^z) \text{ satisfying } v = v^x_{\mathcal{I}}(\theta^x)\}$.

Specifically, $\mathbf{V}^x_{\mathcal{I}}(\eta^z)$ the set of reelection probabilities following policy x that can be generated by a voter best response to $\eta^z \in [0,1]$ (with $\eta^{\neg z} = 0$). To show the desired result we therefore argue that $\mathbf{V}^z_{\mathcal{I}}(\eta^z)$ is decreasing and $\mathbf{V}^{\neg z}_{\mathcal{I}}(\eta^z)$ is increasing.

To argue that $\mathbf{V}^z_{\mathcal{I}}(\eta^z)$ is decreasing, first observe by Lemma 9 and Corollary 43 that $\phi^z(\eta^z) = \min\{\phi^z_-(\eta^z), \phi^z_+(\eta^z)\}$, with $\phi^z_-(\eta^z)$ strictly increasing in η^z and $\phi^z_+(\eta^z)$ strictly decreasing in η^z. Thus, there $\exists$ some $\bar{\eta}^z_z$ where $\phi^z(\eta^z)$ achieves its strict maximum over $[0,1]$, and moreover if $\bar{\eta}^z_z \in (0,1)$ then $\phi^z_-(\eta^z) < (>)(=) \phi^z_+(\eta^z) \iff \eta^z < (>)(=) \bar{\eta}^z_z$.

Suppose first that $c \geq \phi^z(\bar{\eta}^z_z)$. By Lemma 9, if $\eta^z < \bar{\eta}^z_z$ then $\hat{\theta}^z \in \bar{\Theta}^z(\eta^z) \rightarrow \hat{v}^z_\emptyset = 1 > \hat{\rho}^z = 0 \rightarrow \mathbf{V}^z_{\mathcal{I}}(\eta^z) = \{1\}$, and if $\eta^z > \bar{\eta}^z_z$ then $\hat{\theta}^z \in \bar{\Theta}^z(\eta^z) \rightarrow \hat{v}^z_\emptyset = \hat{\rho}^z = 0 \rightarrow \mathbf{V}^z_{\mathcal{I}}(\eta^z) = \{0\}$. $\mathbf{V}^z_{\mathcal{I}}(\eta^z)$ decreasing then immediately follows.

Suppose next that $c < \phi^z(\bar{\eta}^z_z)$. There are three subcases.

(a) If $\eta^z < \bar{\eta}^z_z$ then by Lemma 9 we have $\hat{\theta}^z \in \bar{\Theta}^z(\eta^z) \iff \hat{\theta}^z$ satisfies (i) $\hat{v}^z_\emptyset = \hat{v}^z_z = 1 > \hat{v}^z_{\neg z} = 0$, and (ii) $c > (<) \phi^z_-(\eta^z) \rightarrow \hat{\rho}^z = 1(0)$. Since $\phi^z_-(\eta^z)$ is strictly increasing in η^z, it is easy to see that $\{\rho : \exists \hat{\theta}^z \in \bar{\Theta}^z \text{ with } \rho = \hat{\rho}^z\}$ is an increasing correspondence. Moreover, observe that $v^z_{\mathcal{I}}(\rho^z | \hat{v}^z_\emptyset = \hat{v}^z_z = 1, \hat{v}^z_{\neg z} = 0) = 1 - \rho^z \Pr(\omega \neq x | \mathcal{I})$ is decreasing in ρ^z (that is, more attention to z hurts reelection prospects when the voter's posture is favorable). Thus it immediately follows that $\mathbf{V}^z_{\mathcal{I}}(\eta^z)$ is decreasing over the range $\eta^z < \bar{\eta}^z_z$.

(b) If $\eta^z > \bar{\eta}^z_z$ then by Lemma 9 we have $\hat{\theta}^z \in \bar{\Theta}^z(\eta^z) \iff \hat{\theta}^z$ satisfies (i) $\hat{v}^z_\emptyset = \hat{v}^z_{\neg z} = 0$, (ii) $\phi^z_+(\eta^z) > (<) 0 \rightarrow \hat{v}^z_z = 1(0)$, and (iii) $c > (<) \phi^z_-(\eta^z) \rightarrow \hat{\rho}^z = 1(0)$. Since $\phi^z_+(\eta^z)$ is strictly decreasing in η^z, it is easy to see that both $\{\rho : \exists \hat{\theta}^z \in \bar{\Theta}^z \text{ with } \rho = \hat{\rho}^z\}$ and $\{v : \exists \hat{\theta}^z \in \bar{\Theta}^z \text{ with } v = \hat{v}^z_z\}$ are decreasing correspondences. Moreover, observe that $v^z_{\mathcal{I}}(\rho^z, v^z_z | \hat{v}^z_\emptyset = \hat{v}^z_{\neg z} = 0) = \rho^z v^z_z \cdot \Pr(\omega = z | \mathcal{I})$ is increasing in both v^x_x and ρ^z (that is, more attention to z helps reelection prospects when the voter's posture is adversarial). Thus it immediately follows that $\mathbf{V}^z_{\mathcal{I}}(\eta^z)$ is again decreasing over the range $\eta^z > \bar{\eta}^z_z$.

(c) If η^z is sufficiently close to $\bar{\eta}^z_z$ then by Lemma 9 we have $\hat{\theta}^z \in \bar{\Theta}^z(\eta^z) \rightarrow \hat{\rho}^z = \hat{v}^z_z = 1 > \hat{v}^z_{\neg z} = 0 \rightarrow \mathbf{V}^z_{\mathcal{I}}(\eta^z) = \{\Pr(z = \omega | \mathcal{I})\}$ and constant.

Finally, exactly symmetric arguments show $\mathbf{V}_{\mathcal{I}}^{\neg z}(\eta^z)$ is increasing, beginning again with the observations (by Lemma 9 and Corollary 43) that $\phi^{\neg z}(\eta^z) = \min\{\phi_-^{\neg z}(\eta^z), \phi_+^{\neg z}(\eta^z)\}$, but with $\phi_+^{\neg z}(\eta^z)$ strictly increasing in η^z and $\phi_-^{\neg z}(\eta^z)$ strictly decreasing in η^z. QED

With the preceding lemma in hand, we first prove main text Proposition 10 stating that the incumbent is always truthful when $\pi = \frac{1}{2}$ (i.e., is no ex-ante "popular" policy).

Proof of Proposition 10 Applying Proposition 39 and Lemma 44, to rule out an equilibrium distorted toward a policy $x \in \{A, B\}$ ($\eta^x > 0, \eta^{\neg x}$=0), it suffices to show $\min\{\mathbf{\Delta}_{s=\neg x}^x(0)\} \leq 0$ (intuitively, that there is no electoral benefit to policy x after signal $\neg x$ when the incumbent is believed to be truthful). Given ex-ante policy symmetry and incumbent truthfulness, there always exists a best-response $\hat{\theta}$ in which the voter treats the incumbent identically after either policy, so $\Delta_{s=\neg x}^x(\hat{\theta}) = \rho^x(\Pr(\omega = \neg x|s = x) - \Pr(\omega = x|s = x)) \leq 0$. QED.

We next prove Proposition 11 ruling out "fake leadership" equilibria.

Proof of Proposition 11 Applying Proposition 39 and Lemma 44, to rule out fake leadership equilibria ($\eta^A = 0, \eta^B \in (0, 1)$) it suffices to show that $\min\{\mathbf{\Delta}_{s=A}^B(0)\} \leq 0$ (intuitively, that there is no electoral benefit to the unpopular policy B when the incumbent is believed to be truthful). Recall from the main text that $\bar{\mu}^B < \mu < \bar{\mu}^A < \bar{\mu}_A^A = \bar{\mu}_B^B$.

Suppose first that $\gamma \in (\bar{\mu}^B, \bar{\mu}^A)$ so that $v_\emptyset^A = 1 > v_\emptyset^B = 0$ in a voter best response. Then it is easily verified that $\min\{\mathbf{\Delta}_{s=A}^B(0)\} \leq -(2\Pr(\omega = A|s = A) - 1) \leq 0$.

Suppose next that $\gamma \leq \bar{\mu}^B$, so that the voter's posture is favorable after both policies. Then $\bar{\phi}^B > \bar{\phi}^A$ (by Lemma 40), and there exists some $\hat{\theta} \in \bar{\Theta}(0)$ with $\hat{v}_x^x = \hat{v}^A = 1 > \hat{v}_{\neg x}^x = 0\ \forall x$ and $\hat{\rho}^B \geq \hat{\rho}^A$, so $\Delta_{s=A}^B(\hat{\theta}) =$

$$-\hat{\rho}^A(2\Pr(\omega = A|s = A) - 1) - (\hat{\rho}^B - \hat{\rho}^A)\Pr(\omega = A|s = A) - (1 - \hat{\rho}^B)(1 - \hat{v}^B) \leq 0.$$

Suppose next that $\gamma \in [\bar{\mu}_A^A, \bar{\mu}^A]$ (recalling that $\bar{\mu}_A^A = \bar{\mu}_B^B$) so that the voter has an adversarial posture after both policies. Then $\bar{\phi}^A > \bar{\phi}^B$ (by Lemma 40), and there exists some $\hat{\theta} \in \bar{\Theta}(0)$ with $\hat{v}_x^x = 1 > \hat{v}_{\neg x}^x = \hat{v}^B = 0\ \forall x$ and $\hat{\rho}^A \geq \hat{\rho}^B$, so $\Delta_{s=A}^B(\hat{\theta}) =$

$$-\hat{\rho}^B(2\Pr(\omega = A|s = A) - 1) - (\hat{\rho}^A - \hat{\rho}^B)\Pr(\omega = A|s = A) - (1 - \hat{\rho}^A)\hat{v}^A \leq 0.$$

Finally suppose that $\bar{\mu}_A^A = \bar{\mu}_B^B < \gamma$; then clearly $\mathbf{\Delta}_{s=A}^B(0) = \{0\}$. QED.

We conclude by proving existence and generic uniqueness of sequential equilibrium.

Lemma 45 *A sequential equilibrium of the model exists and is generically unique.*

Proof: It is straightforward to verify from the definitions that for generic model parameters $(\mu, \gamma, \pi, q, c) \in [0,1]^4 \times \mathcal{R}^+$ we have that (i) for any particular fixed $\eta = (\eta^A, \eta^B)$, $\mathbf{\Delta}_{s=B}^A(\eta)$ is a singleton, and (ii) $\mathbf{\Delta}_{s=B}^A(0) \neq \bar{\Delta}_{s=B}^A$.

Suppose first that $\mathbf{\Delta}_{s=B}^A(0) < \bar{\Delta}_{s=B}^A$; then by Proposition 39 there exists a truthful equilibrium. Moreover, by Lemma 44, $\mathbf{\Delta}_{s=B}^A(\eta^A) < \bar{\Delta}_{s=B}^A \; \forall \eta^A > 0$. Hence again by Proposition 39 there cannot exist a pandering equilibrium with $\hat{\eta}^A > 0$.

Suppose next that $\mathbf{\Delta}_{s=B}^A(0) > \bar{\Delta}_{s=B}^A$; then by Proposition 39 there does not exist a truthful equilibrium. In addition, by Lemma 44, $\mathbf{\Delta}_{s=B}^A\left(\eta^A\right)$ is decreasing and satisfies $\mathbf{\Delta}_{s=B}^A(1) \leq 0 < \bar{\Delta}_{s=B}^A \in (0,1)$. Thus, there $\exists$ some $\hat{\eta}^A > 0$ with $\bar{\Delta}_{s=B}^A \in \mathbf{\Delta}_{s=B}^A(\hat{\eta}^A)$, so by Proposition 39 a pandering equilibrium exists at $\hat{\eta}^A$. Moreover, for generic parameters, $\hat{\eta}^A$ must be equal to one of the (at most) four values where $\mathbf{\Delta}_{s=B}^A(\hat{\eta}^A)$ is non-singleton, with $\bar{\Delta}_{s=B}^A \in (\min\{\mathbf{\Delta}_{s=B}^A(\hat{\eta}^A)\}, \max\{\mathbf{\Delta}_{s=B}^A(\hat{\eta}^A)\})$. Thus, by Lemma 44 we have $\mathbf{\Delta}_{s=B}^A(\eta^A) > (<) \bar{\Delta}_{s=B}^A$ for $\eta^A < (>) \hat{\eta}^A$ and no other pandering equilibrium exists. QED.

B.3 Main Proofs

In this appendix we prove Propositions 12—19 and Lemma 14 describing the form of equilibrium across the parameter space. Since fake leadership has been ruled out, for the remaining analysis we return to the notation in the main text, denoting the probability that a low-ability incumbent chooses A after signal B as simply σ (rather than η^A) and assuming throughout that a low-ability incumbent always chooses A after signal A (i.e., $\eta^B = 0$).

Simplified Equilibrium Characterization

We first collect definitions and properties from the main text and preceding Appendices.

With respect to the voter, recall that (i) $\bar{\phi}_-^B > \bar{\phi}_-^A$ and $\bar{\phi}_+^A > \bar{\phi}_+^B$, (ii) $\phi_-^A(\sigma)$ and $\phi_+^B(\sigma)$ are strictly increasing in σ, (iii) $\phi_+^A(\sigma)$ and $\phi_-^B(\sigma)$ are strictly decreasing in

σ, and (iv) $\phi^x_-(\sigma) < (>)(=)\phi^x_+(\sigma) \iff \mu^x_\emptyset(\sigma) > (<)\gamma$, further implying that $\phi^x(\sigma) = \min\{\phi^x_-(\sigma), \phi^x_+(\sigma)\}$.

With respect to the incumbent, having ruled out fake leadership we may focus specifically on incentives after observing the unpopular signal $s = B$. Recall from Appendix A that:

$$\Delta^A_{s=B}(\theta) = \Big((1-\rho^A)\nu^A_\emptyset + \rho^A \Pr(\omega = A|s = B)\Big) - \Big((1-\rho^B)\nu^B_\emptyset + \rho^B \Pr(\omega = B|s = B)\Big)$$

(imposing $1 = \nu^x_x > \nu^x_{\neg x} = 0$ which always holds when $\rho^x > 0$ is a best response), and also

$$\bar{\Delta}^A_{s=B} = \frac{\Pr(\omega = B|s = B) - \Pr(\omega = A|s = B)}{\delta q}.$$

Next recall from the main text that $\hat{q}$ is the unique solution to

$$\delta\hat{q}\left((1-\pi)\hat{q} + \pi(1-\hat{q})\right) = \hat{q} - \pi.$$

It is helpful to observe that $\bar{\Delta}^A_{s=B} < (=) > 1 \iff q < (>) = \hat{q}$. Finally, recall from the main text that $\bar{q}$ is the unique solution to

$$\delta\hat{q}\left((1-\pi)\hat{q} + \pi(1-\hat{q})\right) = \hat{q} - \pi.$$

It is again helpful to observe that $\bar{\Delta}^A_{s=B} < (=) > \Pr(\omega = A|s = B) \iff q < (>) = \bar{q}$.

The simplified equilibrium characterization is then as follows.

Corollary 46 *Profile $(\hat{\sigma}, \hat{\theta})$ is a sequential equilibrium i.f.f. it satisfies Lemma 9 and either*

- $\hat{\sigma} = 0$ *(the incumbent is truthful) and* $\Delta^A_{s=B}(\hat{\theta}) \leq \bar{\Delta}^A_{s=B}$
- $\hat{\sigma} \in (0, 1)$ *(the incumbent panders); and* $\Delta^A_{s=B}(\hat{\theta}) = \bar{\Delta}^A_{s=B}$.

A sequential equilibrium of the model always exists and is generically unique.

Truthful Equilibria

Recall from Proposition 8 that a truthful equilibrium of the CHS model exists i.f.f. either (i) $\gamma \notin (\bar{\mu}^B, \bar{\mu}^A)$ or (ii) $q \geq \hat{q}$. We now provide conditions for existence of a truthful equilibrium in the RA model; Propositions 12 and 13 are then immediate corollaries.

Lemma 47 *There exists a truthful equilibrium of the RA model if and only if:*

- $c \leq \min\{\bar{\phi}^A, \bar{\phi}^B\}$*;*
- $c \in \left(\min\{\bar{\phi}^A, \bar{\phi}^B\}, \max\{\bar{\phi}^A, \bar{\phi}^B\}\right)$ *and* $q \geq \bar{q}$*;*
- $c \geq \max\{\bar{\phi}^A, \bar{\phi}^B\}$ *and either (i)* $\gamma \notin (\bar{\mu}^B, \bar{\mu}^A)$ *or (ii)* $q \geq \hat{q}$*.*

Proof: Suppose first that $c \leq \min\{\bar{\phi}^A, \bar{\phi}^B\}$; then there exists a voter best response $\hat{\theta}$ to truthfulness with full attention ($\hat{\rho}^A = \hat{\rho}^B = 1$), for any such $\hat{\theta}$ we have $\Delta^A_{s=B}(\hat{\theta}) = \Pr(\omega = A|s = B) - \Pr(\omega = B|s = B) < 0 < \bar{\Delta}^A_{s=B}$, so truthfulness is a best response to full attention, and a truthful equilibrium exists.

Suppose next that $c \in \left(\min\{\bar{\phi}^A, \bar{\phi}^B\}, \max\{\bar{\phi}^A, \bar{\phi}^B\}\right)$. Then in any best response $\hat{\theta}$, either $\hat{\rho}^B = 1 > \hat{\rho}^A = 0$ and $\gamma < \bar{\mu}^A$ implying $\hat{\nu}^A = 1$, or $\hat{\rho}^A = 1 > \hat{\rho}^B = 0$ and $\gamma > \bar{\mu}^B$ implying $\hat{\nu}^B = 1$. In either case, $\Delta^A_{s=B}(\hat{\theta}) = \Pr(\omega = A|s = B)$. This in turn is $\leq \bar{\Delta}^A_{s=B}$ (and thus a truthful equilibrium exists) i.f.f. $q \geq \bar{q}$.

Finally suppose that $c \geq \max\{\bar{\phi}^A, \bar{\phi}^B\}$; then there exists a voter best response $\hat{\theta}$ to truthfulness with no attention after either policy, and conditions on the remaining quantities for truthful equilibrium are trivially identical to conditions in the CHS model. QED.

Asymmetric Attention and Pandering Equilibria

The precise structure of equilibrium is relatively complex within the asymmetric attention region when a low-ability incumbent panders. To describe these equilibria first requires a closer examination of how pandering affects the value of attention after each policy.

The Value of Attention with Pandering

Consider two distinct values of attention $\phi^x_s(\sigma)$ and $\phi^{x'}_{s'}(\sigma)$, which are strictly monotonic in σ. It is straightforward to see that their derivatives will have opposite signs, and hence cross at most once over $\sigma \in [0, 1]$, if *either* $x = x'$ *or* $s = s'$. However, single-crossing is not assured when *both* $x \neq x'$ and $s \neq s'$. In particular, in our analysis it will be necessary to compare the value of negative attention $\phi^A_-(\sigma)$ after A and of positive attention $\phi^B_+(\sigma)$, which are both increasing in σ. We thus begin by proving that these functions also cross at most once over $\sigma = [0, 1]$.

Lemma 48 *$\phi_-^A(\sigma)$ and $\phi_+^B(\sigma)$ cross at most once over* $[0,1]$.

Proof: By Lemma 36, the condition $\phi_+^B > (=) \phi_-^A$ can be equivalently written both as $Z(\sigma, \gamma) > (= 0)$, where

$$Z(\sigma; \gamma) = \Pr(y = A) \cdot (\mu - \Pr(y = B | \omega = B)\gamma) - \Pr(y = B) \cdot \Pr(y = A | \omega = B)\gamma,$$

and also $\hat{Z}(\sigma, \gamma) > (= 0)$, where

$$\hat{Z}(\sigma; \gamma) = \Pr(y = A) \cdot \left(\gamma - \frac{\gamma - \mu}{\Pr(y = A | \omega = B)}\right) - \Pr(y = B) \cdot \gamma$$

(Intuitively, $Z(\sigma; \gamma)$ and $\hat{Z}(\sigma; \gamma)$ are distinct functions of σ and γ with potentially distinct derivatives in (σ, γ), but both of whose zeroes over $[0,1]$ are crossings of ϕ_+^B and ϕ_-^A).

Now $Z(\sigma, \gamma)$ is strictly decreasing in γ and $Z(\sigma; \mu) = \Pr(y = A) - \Pr(y = B) > 0$ $\forall \sigma \in [0,1]$; hence, $\phi_+^B - \phi_-^A > 0$ $\forall \sigma \in [0,1]$ when $\gamma \leq \mu$. Next observe that $\hat{Z}(\sigma; \gamma)$ is strictly increasing in σ at any (γ, σ) where both $\gamma > \mu$ and $\hat{Z}(\sigma; \gamma) \geq 0$ (since then $\gamma > \frac{\gamma - \mu}{\Pr(y = A | \omega = B)}$), so $\hat{Z}(\sigma; \gamma)$ and hence also $Z(\sigma; \gamma)$ and $\phi_+^B - \phi_-^A$ satisfy single-crossing in σ. QED

Having shown that single-crossing holds for all necessary pairs for our subsequent analysis, we next introduce several useful definitions.

Definition 49 *For each $(x, s) \in \{A, B\} \times \{-, +\}$, let $\tilde{\phi}_s^s(\sigma)$ denote the unique continuously-differentiable function that extends $\phi_s^x(\sigma)$ linearly over $\mathbb{R}$, and define cutpoints $\sigma_{x,s}^{x',s'}$ and $\sigma_s^x(c)$ as follows:*[1]

- *Let $\sigma_{x,s}^{x',s'}$ denote the unique solution to $\tilde{\phi}_s^x(\sigma) = \tilde{\phi}_{s'}^{x'}(\sigma)$;*
- *Let $\sigma_s^x(c)$ denote the (well-defined) inverse of $\tilde{\phi}_s^x(\sigma)$.*

In short, $\sigma_{x,s}^{x',s'}$ denotes the level of pandering that equates the values of attention $\phi_s^x(\sigma)$ and $\phi_{s'}^{x'}(\sigma)$, while $\sigma_s^x(c)$ denotes the level of pandering that equates the value of attention ϕ_s^x with its exogenous cost c. (We intermittently indicate the dependence of these cutpoints on γ, depending on the context). We now prove several essential properties of these cutpoints.

[1] Specifically, $\tilde{\phi}_s^x(\sigma) = \phi_s^x(\sigma)$ for $\sigma \in [0,1]$, $\left.\frac{\partial \tilde{\phi}_s^x(\sigma)}{\partial \sigma}\right|_{\sigma=0} \cdot \sigma$ for $\sigma < 0$, and $\left.\frac{\tilde{\phi}_s^x(\sigma)}{\partial \sigma}\right|_{\sigma=1} \cdot \sigma$ for $\sigma > 1$.

Lemma 50 *The cutpoints $\sigma_{x,s}^{x',s'}$ satisfy the following:*

- $\mu^x(\sigma_{x-}^{x+}(\gamma)) = \gamma\ \forall x \in \{A, B\}$ *and* $\sigma_N^* = \min\{\max\{\sigma_{A-}^{A+}, 0\}, \max\{\sigma_{B-}^{B+}, 0\}\}$*;*
- $\sigma_{A-}^{B-}(\gamma) \in (0, 1)$ *and is constant in* γ*;*
- $\sigma_{A+}^{B+}(\gamma) \in (0, 1)$ *and is* $< \sigma_{B-}^{B+}$ *when* $\gamma > \mu$*;*
- $\sigma_{A-}^{B-}(\gamma)$ *is strictly increasing in* γ *when* $\sigma_{A-}^{B-}(\gamma) \in [0, 1]$*, and there* $\exists\ \underline{\gamma}, \bar{\gamma}$ *with* $\mu < \underline{\gamma} < \bar{\gamma} < \bar{\mu}^A$ *such that* $\sigma_{A-}^{B+}(\underline{\gamma}) = 0$ *and* $\sigma_{A-}^{B+}(\bar{\gamma}) = \sigma_{A-}^{A+}(\bar{\gamma}) = \sigma_N^*(\bar{\gamma})$.

Proof: The first property is an immediate implication of Lemma 35 and Proposition 8, and the second is easily verified from the definitions.

Proof of third property: We argue that $\gamma > \mu \rightarrow \phi_+^A(\sigma_{B-}^{B+}) < \phi_+^B(\sigma_{B-}^{B+})$; combined with $\phi_+^A(0) < \phi_+^B(0)$ (from Lemma 40), $\phi_+^A(\sigma)$ decreasing in σ and $\phi_+^B(\sigma)$ increasing in σ (from Corollary 43) this yields the desired property. Recall from the main text that there exists a unique level of pandering $\hat{\sigma} \in (0, 1)$ that makes policy choice uninformative and thus satisfies $\mu^A(\hat{\sigma}) = \mu^B(\hat{\sigma}) = \mu$. Further, is easily verified that at $\hat{\sigma}$, for all $x \in \{A, B\}$ we have $\Pr(y = x|\lambda_I = L) = \Pr(y = x|\lambda_I = H) = \Pr(\omega = x)$ (since a high-ability incumbent always chooses correctly). Now suppose that $\mu < \gamma$. Then (i) $\mu^B(\hat{\sigma}) = \mu < \gamma$, (ii) $\mu^B\left(\sigma_{B-}^{B+}\right) = \gamma$, and (iii) $\mu^B(\sigma)$ increasing jointly imply that $\hat{\sigma} < \sigma_{B-}^{B+}$. We last argue that $\phi_+^A(\hat{\sigma}) < \phi_+^B(\hat{\sigma})$, which implies the desired property since $\phi_+^A(\sigma)$ is decreasing and $\phi_+^B(\sigma)$ is increasing. Observe that $\phi_+^A(\hat{\sigma}) < \phi_+^B(\hat{\sigma})$ if and only if

$$\begin{aligned}
& \Pr(\omega = A|y = A)\left(\mu_A^A - \gamma\right) < \Pr(\omega = B|y = B)\left(\mu_B^B - \gamma\right) \\
\iff\quad & \mu^A - \Pr(\omega = A|y = A)\gamma < \mu^B - \Pr(\omega = B|y = B)\gamma \\
\iff\quad & \Pr(\omega = A|y = A) > \Pr(\omega = B|y = B) \\
\iff\quad & \mu \Pr(\omega = A|y = A, \lambda_I = H) + (1 - \mu)\Pr(\omega = A|y = A, \lambda_I = L) \\
>\quad & \mu \Pr(\omega = B|y = B, \lambda_I = H) + (1 - \mu)\Pr(\omega = B|y = B, \lambda_I = L) \\
\iff\quad & \Pr(\omega = A|y = A, \lambda_I = L) > \Pr(\omega = B|y = B, \lambda_I = L) \\
\iff\quad & \frac{\Pr(y = A|\omega = A, \lambda_I = L)\Pr(\omega = A)}{\Pr(y = A|\lambda_I = L)} > \frac{\Pr(y = B|\omega = B, \lambda_I = L)\Pr(\omega = B)}{\Pr(y = B|\lambda_I = L)} \\
\iff\quad & \Pr(y = A|\omega = A, \lambda_I = L) > \Pr(y = B|\omega = B, \lambda_I = L) \\
\iff\quad & q + (1 - q)\sigma > q(1 - \sigma), \text{ which holds } \forall \sigma > 0.
\end{aligned}$$

The first equivalence follows from Lemma 33, the second from $\mu^A(\hat{\sigma}) = \mu^B(\hat{\sigma}) = \mu$, the fourth from $\Pr(\omega = x|y = x, \lambda_I = H) = 1$, and the sixth from $\Pr(y = x|\lambda_I = L) = \Pr(\omega = x)$ at $\hat{\sigma}$. QED.

Proof of fourth property: Recall from the proof of Lemma 48 that $\phi_{+}^{B}(\sigma;\gamma) - \phi_{-}^{A}(\sigma;\gamma) > (=)\, 0$ i.f.f. $Z(\sigma,\gamma) > (=0)$, where $Z(\sigma,\gamma)$ is strictly decreasing in γ and crosses 0 over $\sigma \in [0,1]$ at most once.

We first argue that $\sigma_{A-}^{B+}(\gamma)$ is strictly increasing in γ when $\sigma_{A-}^{B+}(\gamma) \in [0,1]$. For $\gamma < \gamma'$ where both $\sigma_{A-}^{B+}(\gamma) \in [0,1]$ and $\sigma_{A-}^{B+}(\gamma') \in [0,1]$ we have that $Z\left(\sigma_{A-}^{B+}(\gamma);\gamma\right) = 0 \rightarrow Z\left(\sigma_{A-}^{B+}(\gamma);\gamma'\right) < 0$, implying $\sigma_{A-}^{B+}(\gamma')$ such that $\hat{Z}\left(\sigma_{A-}^{B+}(\gamma);\gamma'\right) = 0$ must satisfy $\sigma_{A-}^{B+}(\gamma') > \sigma_{A-}^{B+}(\gamma)$ by single crossing of $Z(\sigma,\gamma)$ over $\sigma \in [0,1]$.

We next argue that there exists a unique $\underline{\gamma} \in \left(\mu, \bar{\mu}^{A}\right)$ that solves $\sigma_{A-}^{B+}\left(\underline{\gamma}\right) = 0$, which is equivalent to $\phi_{+}^{B}\left(0;\underline{\gamma}\right) - \phi_{-}^{A}\left(0;\underline{\gamma}\right) = 0$. To see this, observe that $Z(\sigma;\mu) = \Pr(y=A) - \Pr(y=B) > 0\ \forall \sigma \in [0,1]$ so $\phi_{+}^{B}(0;\mu) > \phi_{-}^{A}(0;\mu)$, and $\phi_{-}^{A}\left(0;\bar{\mu}^{A}\right) = \phi_{+}^{A}\left(0;\bar{\mu}^{A}\right) > \phi_{+}^{B}\left(0,\bar{\mu}^{A}\right)$ (where the equality follows from $\sigma_{A-}^{A+}\left(\bar{\mu}^{A}\right) = 0$ and the inequality from Lemma 35).

Lastly, since $\sigma_{A-}^{B+}(\gamma)$ is strictly increasing in γ, $\sigma_{A-}^{A+}(\gamma)$ is strictly decreasing in γ, $\sigma_{A-}^{B+}\left(\underline{\gamma}\right) = 0 < \sigma_{A-}^{A+}\left(\underline{\gamma}\right)$, and $\sigma_{A-}^{B+}\left(\bar{\mu}^{A}\right) > \sigma_{A-}^{A+}\left(\bar{\mu}^{A}\right) = 0$, there must exist a unique $\bar{\gamma} \in \left(\underline{\gamma}, \bar{\mu}^{A}\right)$ where $\sigma_{A-}^{B+}(\bar{\gamma}) = \sigma_{A-}^{A+}(\bar{\gamma})$. QED.

Having established properties of these critical cutpoints, we are now in a position to *bound* the equilibrium level of pandering σ_{R}^{*} under a variety of different conditions.

Lemma 51 *An equilibrium level of pandering σ_{R}^{*} in the RA model satisfies the following.*

- *If $\gamma < \bar{\gamma}$ then $\sigma_{R}^{*} \leq \sigma_{A-}^{A+}$;*
- *If $\gamma < \underline{\gamma}$ then $\sigma_{R}^{*} < \sigma_{A-}^{B-}$;*
- *If $\gamma \geq \bar{\gamma}$ then $\sigma_{R}^{*} < \sigma_{A+}^{B+}$;*
- *If $\gamma \in [\underline{\gamma}, \bar{\gamma}]$ then $c > (<) \phi_{+}^{B}(\sigma_{A-}^{B+}) = \phi_{-}^{A}(\sigma_{A-}^{B+}) \rightarrow \sigma_{R}^{*} > (<) \sigma_{A-}^{B+}$.*

Proof: We first argue $\gamma \leq \bar{\gamma} \rightarrow \sigma_{R}^{*} \leq \sigma_{A-}^{A+}$. Suppose alternatively that $\sigma_{R}^{*} > \sigma_{A-}^{A+}$; then $\nu^{A} = 0$ in any best response. Supporting such an equilibrium requires that a low-ability incumbent who receives signal B have a strict electoral incentive to choose A; it is easily verified that this in turn requires both that $\nu^{B} < 1$ (so $\sigma_{R}^{*} \leq \sigma_{B-}^{B+}$), and also that $\rho^{A} > \rho^{B}$ (so $\phi^{A}\left(\sigma_{R}^{*}\right) \geq \phi^{B}\left(\sigma_{R}^{*}\right)$). Clearly we cannot have $\gamma \leq \mu$ since then $\sigma_{B-}^{B+} \leq \sigma_{A-}^{A+}$, so suppose instead that $\gamma \in (\mu, \bar{\gamma}]$. Then we have $\sigma_N^* = \sigma_{A-}^{A+}$, $\phi^A\left(\sigma_R^*\right) = \phi_+^A\left(\sigma_R^*\right) < \phi_+^A\left(\sigma_{A-}^{A+}\right) = \phi_-^A\left(\sigma_{A-}^{A+}\right) = \phi_-^A\left(\sigma_N^*\right)$ and $\phi^B\left(\sigma_R^*\right) = \phi_+^B\left(\sigma_R^*\right) > \phi_+^B\left(\sigma_{A-}^{A+}\right) = \phi_+^B\left(\sigma_N^*\right)$. But by the definition of $\bar{\gamma}$ we have $\phi_+^B\left(\sigma_N^*\right) > \phi_-^A\left(\sigma_N^*\right)$ implying $\phi^B\left(\sigma_R^*\right) > \phi^A\left(\sigma_R^*\right)$, a contradiction.

We next argue $\gamma \leq \underline{\gamma} \rightarrow \sigma_R^* < \sigma_{A-}^{B-}$. By the definition of $\underline{\gamma}$ we have we have $\phi_-^A(\sigma) < \phi_+^B(\sigma)\ \forall\sigma$ so $\sigma_{B+}^{B-} < \sigma_{A-}^{B-}$. Thus $\phi_A\left(\sigma_{A-}^{B-}\right) \leq \phi_-^A\left(\sigma_{A-}^{B-}\right) = \phi_-^B\left(\sigma_{A-}^{B-}\right) = \phi^B\left(\sigma_{A-}^{B-}\right)$. Now consider a voter best response $\hat{\theta}$ to σ_{A-}^{B-}. If $c > \phi_-^B\left(\sigma_{A-}^{B-}\right)$ then in any best response, $\nu^B = 1 > \rho^B = 0$; but then $\Delta_{s=B}^A(\hat{\theta}) \leq 0 < \bar{\Delta}_{s=B}^A$ so $\sigma_R^* < \sigma_{A-}^{B-}$. Alternatively, if $c < \phi_-^B\left(\sigma_{A-}^{B-}\right)$ then in any best response $\hat{\theta}$ we have $\rho^B = 1$, and either have $\rho^A = 1$ (if $\phi^A\left(\sigma_{A-}^{B-}\right) = \phi_-^A\left(\sigma_{A-}^{B-}\right) \leq \phi_+^A\left(\sigma_{A-}^{B-}\right)$) or $\rho^A = \nu^A = 0$ (if $\phi^A\left(\sigma_{A-}^{B-}\right) = \phi_+^A\left(\sigma_{A-}^{B-}\right) < \phi_-^A\left(\sigma_{A-}^{B-}\right)$); in either case $\Delta_{s=B}^A(\hat{\theta}) \leq -\left(\Pr(\omega = B|s = B) - \Pr(\omega = A|s = B)\right) < 0 < \bar{\Delta}_{s=B}^A$, so again $\sigma_R^* < \sigma_{A-}^{B-}$.

We next argue that $\gamma \geq \bar{\gamma} \rightarrow \sigma_R^* \leq \sigma_{A+}^{B+}$. By the definition of $\bar{\gamma}$ we have that $\sigma_{A+}^{A-} \leq \sigma_{A+}^{B+} \leq \sigma_{A-}^{B+}$, and further by Lemma 50 we have that $\sigma_{A+}^{B+} \leq \sigma_{B-}^{B+}$. Hence $\phi^A\left(\sigma_{A+}^{B+}\right) = \phi_+^A\left(\sigma_{A+}^{B+}\right) = \phi_+^B\left(\sigma_{A+}^{B+}\right) = \phi^B\left(\sigma_{A+}^{B+}\right)$. We now consider a voter best response $\hat{\theta}$ to σ_{A+}^{B+}. If $c > \phi^A\left(\sigma_{A+}^{B+}\right) = \phi^B\left(\sigma_{A+}^{B+}\right)$, then the voter will replace the incumbent outright after either policy, so $\Delta_{s=B}^A(\hat{\theta}) = 0 < \bar{\Delta}_{s=B}^A$, implying $\sigma_R^* < \sigma_{A-}^{B-}$. Alternatively, if $c < \phi^A\left(\sigma_{A+}^{B+}\right) = \phi^B\left(\sigma_{A+}^{B+}\right)$ then the voter will pay attention after either policy, so $\Delta_{s=B}^A(\hat{\theta}) = -\left(\Pr(\omega = B|s = B) - \Pr(\omega = A|s = B)\right) < 0 < \bar{\Delta}_{s=B}^A$, again implying $\sigma_R^* < \sigma_{A-}^{B-}$.

We last argue that when $\gamma \in \left[\underline{\gamma}, \bar{\gamma}\right]$ we have $\sigma_R^* > (<)\ \sigma_{A-}^{B+}$ when $c > (<)\ \phi_+^B\left(\sigma_{A-}^{B+}\right) = \phi_-^A\left(\sigma_{A-}^{B+}\right)$. Observe that by the definitions of $\underline{\gamma}$ and $\bar{\gamma}$ we have that $\sigma_{A-}^{B+} \leq \sigma_{A+}^{B+} \leq \sigma_{A+}^{A-} < \sigma_{B+}^{B-}$. Hence $\phi^A\left(\sigma_{A-}^{B+}\right) = \phi_-^A\left(\sigma_{A-}^{B+}\right) = \phi_+^B\left(\sigma_{A-}^{B+}\right) = \phi^B\left(\sigma_{A-}^{B+}\right)$. Now consider a voter best response $\hat{\theta}$ to σ_{A-}^{B+}. If $c > \phi^A\left(\sigma_{A-}^{B+}\right) = \phi^B\left(\sigma_{A-}^{B+}\right)$ then the voter will retain the incumbent outright after A and replace her after B, so $\Delta_{s=B}^A(\hat{\theta}) = 1 > \bar{\Delta}_{s=B}^A$, implying $\sigma_R^* > \sigma_{A-}^{B+}$. Alternatively, if $c < \phi^A\left(\sigma_{A-}^{B+}\right) = \phi^B\left(\sigma_{A-}^{B+}\right)$ then the voter will pay attention after either policy, so $\Delta_{s=B}^A(\hat{\theta}) = -\left(\Pr(\omega = B|s = B) - \Pr(\omega = A|s = B)\right) < 0 < \bar{\Delta}_{s=B}^A$, implying $\sigma_R^* < \sigma_{A-}^{B+}$. QED

Finally, we are now in a position to characterize which policy receives more attention in the asymmetric attention region.

Proof of Lemma 14 We first argue that $\gamma < \underline{\gamma} < \bar{\gamma} \rightarrow \phi^B\left(\sigma_R^*\right) > \phi^A\left(\sigma_R^*\right)$, implying $\rho^B \geq \rho^A$. By the definition of $\underline{\gamma}$ we have $\phi_+^B\left(\sigma_R^*\right) > \phi_-^A\left(\sigma_R^*\right)$, and by

Lemma 51 we have $\sigma_R^* \in [0, \sigma_{B-}^{A-})$ which $\rightarrow \phi_-^B(\sigma_R^*) > \phi_-^A(\sigma_R^*)$. Thus $\phi^B(\sigma_R^*) = \min\{\phi_-^B(\sigma_R^*), \phi_+^B(\sigma_R^*)\} > \phi_-^A(\sigma_R^*) \geq \phi^A(\sigma_R^*)$.

We next argue that $\gamma > \bar{\gamma} > \underline{\gamma} \rightarrow \phi^A(\sigma_R^*) > \phi^B(\sigma_R^*)$, implying $\rho^A \geq \rho^B$. By Lemma 51 we have that $\sigma_R^* \in [0, \sigma_{B+}^{A+})$, and by Lemma 50 we have $\sigma_{B+}^{A+} < \sigma_{B+}^{B-}$. Hence $\phi_+^A(\sigma_R^*) > \phi_+^B(\sigma_R^*) = \phi^B(\sigma_R^*)$. Now if $\sigma_R^* \geq \sigma_{A-}^{A+}$ then $\phi^A(\sigma_R^*) = \phi_+^A(\sigma_R^*)$ which yields the desired property, whereas if $\sigma_R^* \leq \sigma_{A-}^{A+} \leq \sigma_N^*$ then $\phi^A(\sigma_R^*) = \phi_-^A(\sigma_R^*) > \phi_+^B(\sigma_R^*)$ from the definition of γ, again yielding the desired property.

We last argue that if $\gamma \in [\underline{\gamma}, \bar{\gamma}]$ we have $c > (<)\, \phi_-^B(\sigma_{A-}^{B+}) = \phi_+^A(\sigma_{A-}^{B+}) \rightarrow \rho^B \leq (\geq)\, \rho^A$. Observe that $\sigma_N^* = \sigma_{A-}^{A+}$, by the definitions of $\underline{\gamma}$ and $\bar{\gamma}$ we have $\sigma_{A-}^{B+} \leq \sigma_{A+}^{B+} \leq \sigma_{A-}^{A+}$, and also $\sigma_{A-}^{A+} < \sigma_{B-}^{B+}$ since $\mu < \underline{\gamma}$. Hence $\forall \sigma \in [0, \sigma_{A-}^{A+}]$ we have $\phi^A(\sigma) = \phi_-^A(\sigma)$ and $\phi^B(\sigma) = \phi_+^B(\sigma)$. Finally by Lemma 51 we have $c > \phi_-^B(\sigma_{A-}^{B+}) \rightarrow \sigma_R^* > \sigma_{A-}^{B+} \rightarrow \phi^A(\sigma_R^*) > \phi^B(\sigma_R^*) \rightarrow \rho^A \geq \rho^B$ and $c < \phi_-^B(\sigma_{A-}^{B+}) \rightarrow \sigma_R^* < \sigma_{A-}^{B+} \rightarrow \phi^A(\sigma_R^*) < \phi^B(\sigma_R^*) \rightarrow \rho^B \geq \rho^A$. QED.

Equilibrium with Moderate-Quality Information

We now use the preceding to fully characterize equilibrium in the asymmetric attention attention region when a low-ability incumbent receives moderate-quality information. Propositions 16 and 18 in the main text are corollaries of this more complete characterization.

Case 1. Suppose that $c \in (\min\{\phi^A(0), \phi^B(0)\}, \max\{\phi^A(0), \phi^B(0)\}]$. Then by Lemma 47, there exists a truthful equilibrium .

Case 2. Suppose that $c \in (\max\{\phi^A(0), \phi^B(0)\}, \max\{\phi^A(\sigma_N^*), \phi^B(\sigma_N^*)\})$. Then $\sigma_N^* \neq 0$ and $\gamma \in (\bar{\mu}^B, \bar{\mu}^A)$. Then in any best response $\hat{\theta}$ to truthfulness we have $\hat{v}^A = 1 > \hat{v}^B = \hat{\rho}^A = \hat{\rho}^B = 0$, implying $\Delta_{s=B}^A(\hat{\theta}) = 1 > \bar{\Delta}_{s=B}^A$, so truthfulness is not a best response to $\hat{\theta}$.

Subcase 2.1: $\gamma \in (\bar{\mu}^B, \underline{\gamma})$. First, since $\phi^A(\sigma) = \phi_-^A(\sigma) < \phi_+^B(\sigma)$ for all $\sigma \in [0, \sigma_N^*]$ (since $\sigma_N^* = \min\{\sigma_{B-}^{B+}, \sigma_{A-}^{A+}\}$) by Lemma 50 the condition reduces to $c \in (\phi_+^B(0), \phi_+^B(\sigma_N^*))$. Thus, there exists a well-defined cutpoint $\sigma_+^B(c) \in (0, \sigma_N^*)$; we argue that there exist an equilibrium with $\hat{\sigma}_R = \sigma_+^B(c)$. First observe that since $\phi_-^A(\sigma) < \phi_+^B(\sigma)\ \forall \sigma \in [0, \sigma_N^*]$, we have that $\hat{v}^A = 1 > \hat{\rho}^A = 0$ is a best response after A. Next observe that since $\sigma_+^B(c) < \sigma_N^* = \min\left\{\sigma_{A-}^{A+}, \sigma_{B-}^{B+}\right\}$, $\hat{\theta}^B$ is a best-response to $\sigma_+^B(c) \iff \hat{\nu}^B = 0$. Since,

$$\Delta_{s=B}^A(\hat{\rho}^B = 0; \hat{\theta}) = 1 > \bar{\Delta}_{s=B}^A > \Delta_{s=B}^A(\hat{\rho}^B = 1; \hat{\theta}) = \Pr(\omega = A | s = B),$$

there exists a best response $\hat{\theta}$ with partial attention $\hat{\rho}^B \in (0,1)$ after B and no attention $\hat{\rho}^A = 0$ after A that supports an equilibrium.

Subcase 2.2: $\gamma \in \left(\underline{\gamma}, \bar{\gamma}\right)$. By Lemma 50 we have $0 < \sigma_{A-}^{B+} < \sigma_{A+}^{B+} < \sigma_{A-}^{A+}$, so the condition reduces to $c \in \left(\phi_-^A(0), \phi_+^B\left(\sigma_{A-}^{A+}\right)\right)$ where $\sigma_{A-}^{A+} = \sigma_N^*$. Thus, there exists a well-defined cutpoint $\min\left\{\sigma_-^A(c), \sigma_+^B(c)\right\} \in \left(0, \sigma_N^*\right)$; we argue that there exists an equilibrium with $\hat{\sigma}_R = \min\left\{\sigma_+^B(c), \sigma_-^A(c)\right\}$.

If $\hat{\sigma}_R = \sigma_+^B(c)$ then $\phi_-^A\left(\sigma_+^B(c)\right) \leq \phi_+^B\left(\sigma_+^B(c)\right) = c$ and $\hat{\theta}^A$ with $\hat{\nu}^A = 1 > \hat{\rho}^A = 0$ is a best response after A. Next observe that since $\sigma_+^B(c) < \sigma_N^* = \min\left\{\sigma_{A-}^{A+}, \sigma_{B-}^{B+}\right\}$, $\hat{\theta}^B$ is a best-response to $\sigma_+^B(c) \iff \hat{\nu}^B = 0$. Since

$$\Delta_{s=B}^A(\hat{\rho}^B = 0; \hat{\theta}) = 1 > \bar{\Delta}_{s=B}^A > \Delta_{s=B}^A(\hat{\rho}^B = 1; \hat{\theta}) = \Pr(\omega = A | s = B),$$

there exists a best response $\hat{\theta}$ with partial attention $\hat{\rho}^B \in (0,1)$ after B and no attention $\hat{\rho}^A = 0$ after A that supports an equilibrium.

If $\hat{\sigma}_R = \sigma_-^A(c)$ then $\phi_+^B\left(\sigma_-^A(c)\right) \leq \phi_-^A\left(\sigma_-^A(c)\right) = c$, and $\hat{\theta}^B$ with $\hat{\rho}^B = \hat{\nu}^B = 0$ is a best response after A. Next, observe that since $\sigma_-^A(c) < \sigma_N^* = \min\left\{\sigma_{A-}^{A+}, \sigma_{B-}^{B+}\right\}$, $\hat{\theta}^A$ is a best response to $\sigma_-^A(c) \iff \hat{\nu}^A = 1$. Since

$$\Delta_{s=B}^A(\hat{\rho}^A = 0; \hat{\theta}) = 1 > \bar{\Delta}_{s=B}^A > \Delta_{s=B}^A(\hat{\rho}^A = 1; \hat{\theta}) = \Pr(\omega = A | s = B),$$

there exists a best response with partial attention $\hat{\rho}^A \in (0,1)$ after A and no attention $\hat{\rho} = 0$ after B that supports an equilibrium.

Subcase 2.3: $\gamma \in \left(\bar{\gamma}, \bar{\mu}^A\right)$. By Lemma 50 we have $0 < \sigma_{A-}^{A+} < \sigma_{A+}^{B+} < \sigma_{A-}^{B+}$, so the condition reduces to $c \in \left(\phi_-^A(0), \phi_-^A\left(\sigma_{A-}^{A+}\right)\right)$ where $\sigma_{A-}^{A+} = \sigma_N^*$. Thus, there exists a well-defined cutpoint $\sigma_-^A(c) \in \left(0, \sigma_N^*\right)$; we argue that there exist an equilibrium with $\hat{\sigma}_R = \sigma_-^A(c)$. First observe that since $\phi_+^B(\sigma) < \phi_-^A(\sigma)\ \forall \sigma \in \left[0, \sigma_N^*\right]$ where $\sigma_N^* = \sigma_{A-}^{A+}$, we have $\hat{\rho}^B = \hat{\nu}^B = 0$ is a best response after B. Next observe that since $\sigma_-^A(c) < \sigma_N^* = \min\left\{\sigma_{A-}^{A+}, \sigma_{B-}^{B+}\right\}$, $\hat{\theta}^A$ is a best-response to $\sigma_-^A(c) \iff \hat{\nu}^A = 1$. Since,

$$\Delta_{s=B}^A(\hat{\rho}^A = 0; \hat{\theta}) = 1 > \bar{\Delta}_{s=B}^A > \Delta_{s=B}^A(\hat{\rho}^A = 1; \hat{\theta}) = \Pr(\omega = A | s = B),$$

there exists a best response $\hat{\theta}$ with partial attention $\hat{\rho}^A \in (0,1)$ after A and no attention $\hat{\rho}^B = 0$ after B that supports an equilibrium. QED

Equilibrium with Low-Quality Information

We last fully characterize equilibria in the asymmetric attention attention region when a low-ability incumbent receives low-quality information ($q \in (\pi, \bar{q})$). Propositions 17 and 19 are corollaries of this more complete characterization.

Recall that $q < \hat{q} \iff \bar{\Delta}^A_{s=B} < \Pr(\omega = A|s = B)$ and

$$c \in \left(\min\left\{\phi^A(0), \phi^B(0)\right\}, \max\left\{\phi^A\left(\sigma^*_N\right), \phi^B\left(\sigma^*_N\right)\right\}\right)$$

We divide up into several cases.

CASE 1: $\gamma \in \left(0, \underline{\gamma}\right)$.

We begin by arguing that (i) $\min\left\{\phi^A(0), \phi^B(0)\right\} = \phi^A_-(0)$ and (ii) $\max\left\{\phi^A\left(\sigma^*_N\right), \phi^B\left(\sigma^*_N\right)\right\} = \phi^B\left(\sigma^*_N\right)$, so that the asymmetric attention condition reduces to

$$c \in \left(\phi^A_-(0), \phi^B\left(\sigma^*_N\right)\right)$$

First observe that $\underline{\gamma} < \bar{\mu}^A \rightarrow \phi^A_-(0) < \phi^A_+(0)$. Second recall from Lemma 40 that $\phi^A_-(0) < \phi^B_-(0)$. Third recall that $\gamma < \underline{\gamma} \rightarrow \phi^A_-(\sigma) < \phi^B_+(\sigma)\ \forall \sigma \in [0,1]$. These immediately yield (i), as well as (ii) when $\gamma \leq \bar{\mu}^B$ so that $\sigma^*_N = 0$. Finally, whenever $\gamma \in \left(\bar{\mu}^B, \bar{\mu}^A\right)$ we have $\phi^B\left(\sigma^*_N\right) = \phi^B_+\left(\sigma^*_N\right)$ and $\phi^A\left(\sigma^*_N\right) = \phi^A_-\left(\sigma^*_N\right)$ which again yields (ii).

We now argue that there exists a pandering equilibrium at

$$\hat{\sigma}_R = \min\{\sigma^B_-(c), \sigma^A_-(c), \sigma^{A+}_{A-}\}.$$

To do so observe that $\gamma < \bar{\mu}^A \rightarrow \sigma^{A+}_{A-} \in (0,1)$ and σ^{B-}_{A-} is constant in γ. We now examine three exhaustive and mutually exclusive conditions on the cost of attention c.

Subcase 1.1 (High Attention). $c \in \left(\phi^A_-(0), \phi^A_-\left(\min\left\{\sigma^{A+}_{A-}, \sigma^{B-}_{A-}\right\}\right)\right)$. It is easily verified that $0 < \sigma^A_-(c) < \min\left\{\sigma^B_-(c), \sigma^{A+}_{A-}\right\}$ so $\hat{\sigma}_R = \sigma^A_-(c)$. Clearly, any $\hat{\theta}^A$ such that $\hat{\nu}^A = 1$ is a best response to $\sigma^A_-(c)$. Next we have $c = \phi^A_-\left(\sigma^A_-(c)\right)$ and

$\phi_-^A\left(\sigma_-^A(c)\right) < \phi_-^B\left(\sigma_-^A(c)\right)$ and $\phi_-^A\left(\sigma_-^A(c)\right) < \phi_+^B\left(\sigma_-^A(c)\right)$, so any $\hat{\theta}^B$ that is a best response to $\sigma_-^A(c)$ must have $\hat{\rho}^B = 1$. Thus, we have:

$$\begin{aligned}\Delta_{s=B}^A(\hat{\rho}^A = 0; \hat{\theta}) &= \Pr(\omega = A|s = B) > \bar{\Delta}_{s=B}^A > \Delta_{s=B}^A(\hat{\rho}^A = 1; \hat{\theta}) \\ &= -\left(\Pr(\omega = B|s = B) - \Pr(\omega = A|s = B)\right),\end{aligned}$$

and there exists a best response to $\sigma_-^A(c)$ with partial attention $\hat{\rho}^A \in (0,1)$ and a favorable posture $\hat{\nu}^A = 1$ after A, and full attention $\hat{\rho}^B = 1$ after B.

Subcase 1.2 (Medium Attention). $c \in \left(\phi_-^A\left(\min\left\{\sigma_{A-}^{A+}, \sigma_{A-}^{B-}\right\}\right), \phi^B\left(\min\left\{\sigma_{A-}^{A+}, \sigma_{A-}^{B-}\right\}\right)\right)$.

We first argue that for this case to hold, γ must be such that $\sigma_{A-}^{A+} < \sigma_{A-}^{B-}$. First recall that by Lemma 50 that $\phi_+^B(\sigma) > \phi_-^A(\sigma)\ \forall\sigma$ when $\gamma < \underline{\gamma}$, which $\rightarrow \sigma_{B-}^{B+} < \sigma_{B-}^{A-}$. Next, if instead we had $\sigma_{A-}^{B-} \leq \sigma_{A-}^{A+}$ then the interval would reduce to $\left(\phi_-^A\left(\sigma_{A-}^{B-}\right), \phi_-^B\left(\sigma_{A-}^{B-}\right)\right)$ which is empty. Concluding, this case may be simplified to $\sigma_{A-}^{A+} < \sigma_{A-}^{B-}$ and

$$c \in \left(\phi_-^A\left(\sigma_{A-}^{A+}\right), \phi^B\left(\sigma_{A-}^{A+}\right)\right).$$

It is easily verified that $\sigma_{A-}^{A+} < \min\left\{\sigma_-^B(c), \sigma_-^A(c)\right\}$ so $\hat{\sigma}_R = \sigma_{A-}^{A+}$.

Now clearly any $\hat{\theta}^A$ with $\hat{\rho}^A = 0$ is a best response to σ_{A-}^{A+}, and any $\hat{\theta}^B$ with $\hat{\rho}^B = 1$ is a best response to σ_{A-}^{A+}. Thus, we have that

$$\Delta_{s=B}^A(\hat{\nu}^A = 1; \hat{\theta}) = \Pr(\omega = A|s = B) > \bar{\Delta}_{s=B}^A > \Delta_{s=B}^A(\hat{\nu}^A = 0; \hat{\theta}) = -\Pr(\omega = B|s = B),$$

and there exists a best response to σ_{A-}^{A+} with no attention $\hat{\rho}^A = 0$ and a mixed posture $\hat{\nu}^A \in (0,1)$ after A, and full attention $\hat{\rho}^B = 1$ after B.

Subcase 1.3 (Low Attention). $c \in \left(\phi^B\left(\min\left\{\sigma_{A-}^{A+}, \sigma_{A-}^{B-}\right\}\right), \phi^B\left(\sigma_N^*\right)\right)$.

We first argue that this case may be simplified to $\gamma < \mu$ and

$$c \in \left(\phi_-^B\left(\min\left\{\sigma_{A-}^{A+}, \sigma_{A-}^{B-}\right\}\right), \phi_-^B\left(\max\left\{\sigma_{B-}^{B+}, 0\right\}\right)\right).$$

To see this, first observe that when $\gamma = \mu$ we have $\sigma_N^* = \sigma_{A-}^{A+} = \sigma_{B-}^{B+}$, so $\phi_-^B\left(\sigma_N^*\right) = \phi_+^B\left(\sigma_N^*\right) > \phi_-^A\left(\sigma_N^*\right)$ (from $\mu < \bar{\gamma}$) implying $\sigma_{B-}^{B+} = \sigma_{A-}^{A+} < \sigma_{A-}^{B-}$. Next since σ_{B-}^{B+} is increasing in γ, σ_{A-}^{A+} is decreasing in γ, and σ_{A-}^{B-} is constant in γ (by Lemma 50), we have that $\sigma_{A-}^{A+} < \sigma_{A-}^{B-}$ for $\gamma \in [\mu, \bar{\gamma}]$ and $\sigma_{B-}^{B+} < \sigma_{A-}^{B-}$ for $\gamma < \mu$. Consequently, the condition reduces to $c \in \left(\phi^B\left(\sigma_{A-}^{A+}\right), \phi^B\left(\sigma_{A-}^{A+}\right)\right)$ when $\gamma \in [\mu, \bar{\gamma})$ (which is empty) and $c \in \left(\phi_-^B\left(\min\left\{\sigma_{A-}^{A+}, \sigma_{A-}^{B-}\right\}\right), \phi_-^B\left(\max\left\{\sigma_{B-}^{B+}, 0\right\}\right)\right)$ when $\gamma < \mu$, which is always nonempty since $\phi_-^B(\sigma)$ is decreasing in σ and $\sigma_{B-}^{B+} < \min\left\{\sigma_{A-}^{A+}, \sigma_{A-}^{B-}\right\}$.

Next, it is easily verified that $0 < \sigma_{-}^{B}(c) < \sigma_{A-}^{A+} < \sigma_{-}^{A}(c)$ so $\hat{\sigma}_R = \sigma_{-}^{B}(c)$. Clearly, any $\hat{\theta}^B$ such that $\hat{v}^B = 1$ is a best response to $\sigma_{-}^{B}(c)$. Next, $\phi^A\left(\sigma_{-}^{B}(c)\right) = \phi_{-}^{A}\left(\sigma_{-}^{B}(c)\right)$ (by $\sigma_{-}^{B}(c) < \sigma_{A-}^{A+}$), which is $< \phi_{-}^{B}\left(\sigma_{-}^{B}(c)\right)$ (by $\sigma_{-}^{B}(c) < \sigma_{A-}^{B-}$) which is $= c$, so $\hat{\theta}^A$ is a best response to $\sigma_{-}^{B}(c)$ i.f.f. $\hat{v}^A = 1 > \hat{\rho}^A = 0$. Thus, we have that:

$$\Delta_{s=B}^{A}(\hat{\rho}^B = 1; \hat{\theta}) = \Pr(\omega = A|s = B) > \bar{\Delta}_{s=B}^{A} > \Delta_{s=B}^{A}(\hat{\rho}^B = 0; \hat{\theta}) = 0,$$

so there exists a best response to $\sigma_{-}^{B}(c)$ with partial attention $\hat{\rho}^B \in (0,1)$ and a favorable posture $\hat{v}^B = 1$ after B, and no attention $\hat{\rho}^A = 0$ with a favorable posture $\hat{v}^A = 1$ after A.

CASE 2: $\gamma \in \left[\underline{\gamma}, \bar{\gamma}\right]$

We begin by recalling useful observations from Lemma 50: (i) $\mu < \underline{\gamma} < \gamma \to \sigma_N^* = \max\left\{0, \sigma_{A+}^{A-}\right\} < \sigma_{B+}^{B-}$ and also $\phi^x(\sigma) = \phi_+^x(\sigma)\ \forall \sigma \in \left[0, \sigma_N^*\right]$, (ii) $\sigma_{B+}^{A-} \in (0, \sigma_N^*)$, and (iii) $\phi_+^A(0) > \phi_+^B(0)$ (and so $\sigma_{B+}^{A+} \in (0,1)$). Combining these observations yields that the cost condition reduces to

$$c \in \left(\phi_+^B(0), \phi_+^B\left(\sigma_N^*\right)\right)$$

From these properties it is also easily verified that $0 < \sigma_{B+}^{A-} < \phi_{B+}^{A+} < \sigma_{A+}^{A-} < \phi_{B+}^{B-}$.

We now argue that there exists a pandering equilibrium at

$$\hat{\sigma}_R = \min\left\{\max\left\{\sigma_+^B(c), \sigma_-^A(c)\right\}, \sigma_{A-}^{A+}\right\}$$

. To do we examine three exhaustive mutually exclusive conditions on the cost.

Subcase 2.1 (High attention favoring A): $c \in \left(\phi_+^B(0), \phi_+^B\left(\sigma_{B+}^{A-}\right)\right)$

It is easily verified that $\sigma_-^A(c) < \sigma_+^B(c) < \sigma_{A-}^{A+} < \phi_{B-}^{B+}$; we argue that there exists an equilibrium with $\hat{\sigma}_R = \sigma_+^B(c)$. Using this we have that $\hat{\theta}^A$ is a best response after A i.f.f. $\hat{v}_A = \hat{\rho}^A = 1$ and $\hat{\theta}^B$ is a best response after B i.f.f. $\hat{v}_B = 0$. Thus, we have that:

$$\begin{aligned}\Delta_{s=B}^{A}(\hat{\rho}^B = 0; \hat{\theta}) &= \Pr(\omega = A|s = B) > \bar{\Delta}_{s=B}^{A} > \Delta_{s=B}^{A}(\hat{\rho}^B = 1; \hat{\theta}) \\ &= -\left(\Pr(\omega = B|s = B) - \Pr(\omega = A|s = B)\right),\end{aligned}$$

so there exists a best response to $\sigma_+^B(c)$ with partial attention $\hat{\rho}^B \in (0,1)$ and an adversarial posture $\hat{v}^B = 0$ after B, and full attention $\hat{\rho}^A = 1$ after A.

Subcase 2.2 (High attention favoring B): $c \in (\phi_+^B(\sigma_{B+}^{A-}), \phi_-^A(\sigma_{A-}^{A+}))$

It is easily verified that $\sigma_+^B(c) < \sigma_-^A(c) < \sigma_{A-}^{A+}$; we argue that there exists an equilibrium with $\hat{\sigma}_R = \sigma_-^A(c)$. Using this we have that $\hat{\theta}^A$ is a best response after A i.f.f. $\hat{\nu}_A = 1$ and $\hat{\theta}^B$ is a best response after B i.f.f. $\hat{\nu}_B = 0 < \hat{\rho}_B = 1$. Thus, we have:

$$\begin{aligned}\Delta_{s=B}^A(\hat{\rho}^A = 0; \hat{\theta}) &= \Pr(\omega = A|s = B) > \bar{\Delta}_{s=B}^A > \Delta_{s=B}^A(\hat{\rho}^A = 1; \hat{\theta}) \\ &= -(\Pr(\omega = B|s = B) - \Pr(\omega = A|s = B)),\end{aligned}$$

and there exists a best response to $\sigma_-^A(c)$ with partial attention $\hat{\rho}^A \in (0,1)$ and a favorable posture $\hat{\nu}^A = 1$ after A, and full attention $\hat{\rho}^B = 1$ after B.

Subcase 2.3 (Medium attention): $c \in (\phi_-^A(\sigma_{A-}^{A+}), \phi_+^B(\sigma_{A-}^{A+}))$

It is easily verified that $\sigma_+^B(c) < \sigma_{A-}^{A+} < \sigma_-^A(c)$; we argue that there exists an equilibrium with $\hat{\sigma}_R = \sigma_{A-}^{A+}$. Using this we have that $\hat{\theta}^A$ is a best response after A i.f.f. $\hat{\rho}_A = 0$ and that every $\hat{\theta}^B$ that is a best response after B satisfies $\hat{\rho}^B = 1$. Thus, we have that

$$\Delta_{s=B}^A(\hat{\nu}^A = 1; \hat{\theta}) = \Pr(\omega = A|s = B) > \bar{\Delta}_{s=B}^A > \Delta_{s=B}^A(\hat{\nu}^A = 0; \hat{\theta}) = -\Pr(\omega = B|s = B),$$

and there exists a best response to σ_{A-}^{A+} with no attention $\hat{\rho}^A = 0$ and a mixed posture $\hat{\nu}^A \in (0,1)$ after A, and full attention $\hat{\rho}^B = 1$ after B.

CASE 3: $\gamma \in (\bar{\gamma}, 1]$

We begin by recalling useful observations from Lemma 50: (i) $\mu < \bar{\gamma} < \gamma \to \sigma_N^* = \max\{0, \sigma_{A+}^{A-}\} < \sigma_{B+}^{B-}$, (ii) $\phi^x(\sigma) = \phi_+^x(\sigma)\ \forall \sigma \in [0, \sigma_N^*]$, (iii) $\phi_+^B(\sigma) < \phi_-^A(\sigma)$ for $\sigma \in [0, \sigma_N^*]$, and (iv) $\phi_+^A(0) > \phi_+^B(0)$ (and so $\sigma_{B+}^{A+} \in (0,1)$), and (v) $0 < \sigma_{B+}^{A+} < \sigma_{B+}^{B-}$. Combining these observation yields that the cost condition reduces to:

$$c \in (\phi_+^B(0), \phi_+^A(\sigma_N^*)).$$

From these properties it is also easily verified that $\sigma_{A-}^{A+} < \sigma_{B+}^{A+} < \sigma_{A-}^{B+}$. We now argue that there exists a pandering equilibrium at

$$\hat{\sigma}_R = \min\{\sigma_+^B(c), \sigma_+^A(c)\}.$$

To do we examine two exhaustive and mutually exclusive conditions on the cost c.

Subcase 3.1 (High attention): $c \in \left(\phi_+^B(0), \phi_+^B\left(\phi_{B+}^{A+}\right)\right)$

It is straightforward that $\sigma_+^B(c) < \sigma_+^A(c)$; we argue that there exists an equilibrium with $\hat{\sigma}_R = \sigma_+^B(c)$. Since $\sigma_+^B(c) < \sigma_{B+}^{A+} < \sigma_{B-}^{B+}$ we have that $\hat{\theta}^B$ is a best response to $\sigma_+^B(c)$ if and only if $\hat{v}^B = 0$. Next we argue that $c < \min\left\{\phi_+^A\left(\sigma_+^B(c)\right), \phi_-^A\left(\sigma_+^B(c)\right)\right\}$ so that in any best response $\hat{\theta}^A$ to $\sigma_+^B(c)$ we must have $\hat{\rho}^A = 1$. To see this, observe that (a) $\gamma > \bar{\gamma} \rightarrow \phi_+^B(\sigma) < \phi_-^A(\sigma) > \forall\sigma \in [0,1]$ (by Lemma 50) so $c = \phi_+^B\left(\sigma_+^B(c)\right) < \phi_-^A\left(\sigma_+^B(c)\right)$, and (b) $c = \phi_+^B\left(\sigma_+^B(c)\right) < \phi_+^B\left(\sigma_{B+}^{A+}\right) < \phi_+^A\left(\sigma_{B+}^{A+}\right) < \phi_+^A\left(\sigma_+^B(c)\right)$.

Thus, we have that:

$$\begin{aligned}\Delta_{s=B}^A(\hat{\rho}^B = 0; \hat{\theta}) &= \Pr(\omega = A | s = B) > \bar{\Delta}_{s=B}^A > \Delta_{s=B}^A(\hat{\rho}^B = 1; \hat{\theta}) \\ &= -\left(\Pr(\omega = B | s = B) - \Pr(\omega = A | s = B)\right),\end{aligned}$$

so there exists a best response to $\sigma_+^B(c)$ with partial attention $\hat{\rho}^B \in (0,1)$ and an adversarial posture $\hat{v}^B = 0$ after B, and full attention $\hat{\rho}^A = 1$ after A.

Subcase 3.2 (Low attention): $c \in \left(\phi_+^A\left(\sigma_{B+}^{A+}\right), \phi_+^A\left(\sigma_{A-}^{A+}\right)\right)$

It is straightforward to see that $\sigma_+^A(c) < \sigma_+^B(c)$; we argue that there exists an equilibrium with $\hat{\sigma}_R = \sigma_+^A(c)$. Since $\sigma_+^A(c) \in \left(\sigma_{A-}^{A+}, \sigma_{B+}^{A+}\right)$, we have that $\hat{\theta}^A$ is a best response to $\sigma_+^A(c)$ if and only if $\hat{v}^A = 0$. Next, since $\sigma_+^A(c) < \sigma_{B+}^{A+} < \sigma_{B+}^{B-}$ we have that $c = \phi_+^A\left(\sigma_+^A(c)\right) > \phi_+^B\left(\sigma_+^A(c)\right) = \phi^B\left(\sigma_+^A(c)\right)$, so that $\hat{\theta}^B$ is a best response to $\sigma_+^A(c)$ if and only if $\hat{v}^B = \hat{\rho}^B = 0$.

Thus, we have that:

$$\Delta_{s=B}^A(\hat{\rho}^A = 1; \hat{\theta}) = \Pr(\omega = A | s = B) > \bar{\Delta}_{s=B}^A > \Delta_{s=B}^A(\hat{\rho}^A = 0; \hat{\theta}) = 0,$$

so there exists a best response to $\sigma_+^A(c)$ with partial attention $\hat{\rho}^A \in (0,1)$ and an adversarial posture $\hat{v}^A = 0$ after A, and no attention $\hat{\rho}^B = 0$ and an adversarial posture $\hat{v}^B = 0$ after B.

B.4 Voter Welfare

In this Appendix we prove main text results about voter welfare.

Proof of Lemma 20 There are three parts of the utility difference between the two models, where the last term represents the net loss in accountability in the first period. To see this, we can write the first period voter expected utilities in equilibria for each model as follows:

$$\Pr(\lambda_I = H) + \Pr(\lambda_I = L)\Bigg(\Pr(\omega = A)(\Pr(s = A|\omega = A)+$$

$$\Pr(s = B|\omega = A)\sigma^*) + \Pr(\omega = B)\Pr(s = B|\omega = B)(1 - \sigma^*)\Bigg) =$$

$$\mu + (1 - \mu)\Bigg(\pi(q + (1 - q)\sigma^*) + (1 - \pi)q(1 - \sigma^*)\Bigg),$$

where σ^* is the equilibrium pandering level for each model. Now if we take the difference of these values between the two models and simplify the expression, we get

$$-(1 - \mu)(q - \pi)(\sigma^*_R - \sigma^*_N).$$

As for the first two terms, they represent the second period benefit of paying attention. Let h^R and h^N denote the probability that a high-ability incumbent will be reelected in each model. For general value of h, the second period expected benefit equals to

$$\delta(h + (1 - h)q).$$

Therefore, second period net benefit (excluding the cost of paying attention) in the rational attention model is

$$\delta(h^R + (1 - h^R)q) - \delta(h^N + (1 - h^N)q) = \delta(1 - q)(h^R - h^N).$$

Now we need to calculate $\delta(1 - q)(h^R - h^N)$. There are several cases to consider.

High Attention ($\rho^x > 0\ \forall x$): If attention is at least sometimes acquired after either policy then $\phi^x = \min\{\phi^x_-, \phi^x_+\} \geq c\ \forall x$. In the rational attention model expected utility can therefore be calculated "as if" the voter was always pays attention, so

$$h^R = \Pr(y = A)(\Pr(\omega = A|y = A)\mu^A_A + \Pr(\omega = B|y = A)\gamma)+$$
$$\Pr(y = B)(\Pr(\omega = B|y = B)\mu^B_B + \Pr(\omega = A|y = B)\gamma).$$

As for h^N there are two cases:

- If $\gamma < \mu$ (the incumbent is strong) then in the CHS equilibrium $\nu^x > 0\ \forall x$, so expected utility can be calculated "as if" the incumbent is always reelected and

$$h^N = \mu = \Pr(y = A)(\Pr(\omega = A|y = A)\mu^A_A + \Pr(\omega = B|y = A)\mu^B_A)+$$
$$\Pr(y = B)(\Pr(\omega = B|y = B)\mu^B_B + \Pr(\omega = A|y = B)\mu^A_B),$$

where the quantities in the decomposition that depend on the incumbent's strategy are calculated using the equilibrium pandering level σ_R^* in the *rational attention* model. Therefore the anticipated net benefit of attention is:

$$\delta(1-q)(h^R - h^N) - c = \Pr(y = A)(\delta(1-q)\Pr(\omega = B|y = A)(\gamma - \mu_A^B) - c) + \\ \Pr(y = B)(\delta(1-q)\Pr(\omega = A|y = B)(\gamma - \mu_B^A) - c) = \\ \Pr(y = A)(\phi_-^A - c) + \Pr(y = B)(\phi_-^B - c).$$

- If $\gamma > \mu$ (the incumbent is weak), then in the CHS equilibrium $\nu^x < 1 \; \forall x$, so expected utility may be calculated "as if" the incumbent is never reelected, and

$$h^N = \gamma = \Pr(y = A)(\Pr(\omega = A|y = A)\gamma + \Pr(\omega = B|y = A)\gamma) + \\ \Pr(y = B)(\Pr(\omega = B|y = B)\gamma + \Pr(\omega = A|y = B)\gamma),$$

where again the quantities in the decomposition are calculated using σ_R^*. Therefore the anticipated net benefit of information is:

$$\delta(1-q)(h^R - h^N) - c = \Pr(y = A)(\delta(1-q)\Pr(\omega = A|y = A)(\mu_A^A - \gamma) - c) + \\ \Pr(y = B)(\delta(1-q)\Pr(\omega = B|y = B)(\mu_B^B - \gamma) - c) = \\ \Pr(y = A)(\phi_+^A - c) + \Pr(y = B)(\phi_+^B - c).$$

Medium Attention ($\rho^A = 1 > \rho^A = 0 \;\; \forall x$): In the rational attention equilibrium the voter always pays attention after policy B but never pays attention after policy A and is indifferent between incumbent and challenger.

- If $\gamma < \mu$ (the incumbent is strong) we can calculate expected utility in the rational attention model as if the voter never acquires information and always retains the incumbent after policy A, so

$$h^R = \Pr(y = A)(\Pr(\omega = A|y = A)\mu_A^A + \Pr(\omega = B|y = A)\mu_A^B) + \\ \Pr(y = B)(\Pr(\omega = B|y = B)\mu_B^B + \Pr(\omega = A|y = B)\gamma),$$

and the overall second period net benefit of information is

$$\delta(1-q)(h^R - h^N) - P(y = B)c = \Pr(y = B)(\delta(1-q)\Pr(\omega = A|y = B)(\gamma - \mu_B^A) - c) = \\ \Pr(y = B)(\phi_-^B - c).$$

- If $\gamma > \mu$ (the incumbent is weak) we can calculate expected utility in the rational attention model as if the voter never pays attention and always replaces the incumbent after policy A, so

$$h^R = \Pr(y = A)(\Pr(\omega = A|y = A)\gamma + \Pr(\omega = B|y = A)\gamma) + \\ \Pr(y = B)(\Pr(\omega = B|y = B)\mu_B^B + \Pr(\omega = A|y = B)\gamma),$$

and the overall second period net benefit of information is

$$\delta(1-q)(h^R - h^N) - P(y = B)c = \Pr(y = B)(\delta(1-q)\Pr(\omega = A|y = B)(\mu_B^B - \gamma) - c) = \\ \Pr(y = B)(\phi_+^B - c).$$

Observe that in this case, for Rational attention model we have $\phi^A = \min\{\phi_-^A, \phi_+^A\} < c$.

Low Attention ($\rho^x < 1\ \forall x$) In the rational attention equilibrium the voter at least sometimes chooses *not* to acquire information after either policy. In addition, it is easily verified that in all low-attention regions we have $\nu^x > 0\ \forall x$ if the incumbent is strong ($\gamma < \mu$) and $\nu^x < 1\ \forall x$ if the incumbent is weak ($\gamma > \mu$). Hence, expected utility in the rational attention model can be calculated as if the voter never pays attention, always retains a strong incumbent, and never retains a weak incumbent. In the CHS model expected utility can also be calculated as if the voter always retains a strong incumbent and never retains a weak incumbent, so there is no anticipated net benefit of attention. Further in the RA model we have $\phi^x = \min\{\phi_-^x, \phi_+^x\} \leq c\ \forall x$.[2]
QED

Proof of Proposition 21 When a low-ability incumbent receives moderate-quality information we have $\sigma_R^* \leq \sigma_N^*$, so

$$U_V^R - U_V^N = \underbrace{\Pr(y = A) \cdot \max\left\{\phi_s^A - c, 0\right\}}_{\geq 0} + \underbrace{\Pr(y = B) \cdot \max\left\{\phi_s^B - c, 0\right\}}_{\geq 0} \\ - \underbrace{(1-\mu)}_{>0}\underbrace{(q-\pi)}_{>0}\underbrace{\left(\sigma_R^* - \sigma_N^*\right)}_{\leq 0} \geq 0.$$

Note that when the information is sometimes acquired after at least one policy choice, $\sigma_R^* < \sigma_N^*$ so the third term becomes strictly positive and rational attention

[2]Note that there still might be overall change in the expected utility due to the first period utility through different pandering levels.

strictly increases the expected utility of the voter. Alternatively, when the voter never pays attention, $\sigma_R^* = \sigma_N^*$ and the whole expression equals to 0 (the voter cannot be strictly better off if the information is never acquired). QED.

Proof of Proposition 22 We directly consider the case of $\gamma < \mu$. The case of $\gamma \in (\bar{\gamma}, \bar{\mu}_x^x)$ is shown with symmetric but slightly simplified arguments; for the remaining cases values of γ it is straightforward to verify that $\sigma_R^* \leq \sigma_N^*$ so the voter is at least weakly better off in the rational attention model.

If $c > \phi^B(\sigma_N^*)$ the voter never pays attention, equilibrium of the two models is identical, and so the voter's utility is the same in both models.

If $c < \phi_-^A(0)$ the incumbent is truthful in both models, so there is no accountability cost. From the equilibrium characterization we generically have $\rho^x = 1 \implies \phi^x - c > 0 \ \forall x$, so

$$\begin{aligned} U_V^R - U_V^N &= \underbrace{\Pr(y = A) \cdot \max\left\{\phi_-^A - c, 0\right\}}_{>0} + \underbrace{\Pr(y = B) \cdot \max\left\{\phi_-^B - c, 0\right\}}_{>0} \\ &\quad - \underbrace{(1 - \mu)}_{>0} \underbrace{(q - \pi)}_{>0} \underbrace{\left(\sigma_R^* - \sigma_N^*\right)}_{=0} > 0, \end{aligned}$$

and the voter is strictly better off in the rational attention model.

If $c \in (\max\{\phi_-^B(\sigma_{A-}^{B-}), \phi_-^B(\sigma_{A+}^{A-})\}, \phi^B(\sigma_N^*))$ it is easily verified from the equilibrium characterization that $\sigma_R^* > \sigma_N^*$ (either $\sigma_R^* > 0 = \sigma_N^*$ or $\sigma_R^* > \sigma_{B-}^{B+} = \sigma_N^*$). Thus, the accountability cost is strictly positive. Moreover, from construction of the equilibrium we have $\rho^x < 1 \rightarrow \phi^x(\sigma_R^*) - c \leq 0$ and $\phi^x(\sigma_R^*) = \phi_-^x(\sigma_R^*) \ \forall x$ so

$$\begin{aligned} U_V^R - U_V^N &= \underbrace{\Pr(y = A) \cdot \max\left\{\phi_-^A - c, 0\right\}}_{=0} + \underbrace{\Pr(y = B) \cdot \max\left\{\phi_-^B - c, 0\right\}}_{=0} \\ &\quad - \underbrace{(1 - \mu)}_{>0} \underbrace{(q - \pi)}_{>0} \underbrace{\left(\sigma_R^* - \sigma_N^*\right)}_{>0} < 0. \end{aligned}$$

Finally, if $c \in (\phi_-^A(0), \max\{\phi_-^B(\sigma_{A-}^{B-}), \phi_-^B(\sigma_{A+}^{A-})\})$ we show that there is a unique cost cutoff $\hat{c}(\gamma)$. The equilibrium level of pandering in the rational attention model is $\sigma_R^* = \min\{\sigma^*, \sigma_{A-}^{A+}\}$ where $\phi_-^A(\sigma^*) = c$. Since ϕ_-^A is an increasing function in σ we always have $\phi_-^A(\sigma_R^*) <= c$. Moreover σ_R^* is weakly increasing in c and ϕ_-^B is strictly decreasing in σ, $\Pr(y = B)$ is strictly decreasing in σ and therefore it is weakly decreasing in c (σ_R^* is weakly increasing in c). Overall, when c increases:

$$U_V^R - U_V^N = \underbrace{\Pr(y=A)\cdot\max\left\{\phi_-^A - c, 0\right\}}_{=0} + \underbrace{\underbrace{\Pr(y=B)}_{\text{weakly decreasing}} \cdot\max\left\{\underbrace{\phi_-^B}_{\text{weakly decreasing}} - \underbrace{c}_{\text{strictly increasing}}, 0\right\}}_{\text{weakly decreasing}}$$
$$- \underbrace{(1-\mu)}_{>0}\underbrace{(q-\pi)}_{>0}\underbrace{\left(\sigma_R^* - \sigma_N^*\right)}_{\text{weakly increasing}}.$$

Meaning $U_V^R - U_V^N$ is weakly decreasing in c. Now we show that this expected utility difference is also strictly decreasing in c. For this, we only need to account for the region where σ_R^* is constant in c[3] i.e, when $c \in (\phi_-^A(\sigma_{A+}^{A-}), \phi_-^B(\sigma_{A+}^{A-}))$. For these cost levels, in equilibrium of the rational attention model we have $\sigma_R^* = \sigma_{A+}^{A-}$ and $c < \phi_-^B(\sigma_R^* = \sigma_{A+}^{A-})$. Overall, we have

$$U_V^R - U_V^N = \underbrace{\Pr(y=A)\cdot\max\left\{\phi_-^A - c, 0\right\}}_{=0} + \underbrace{\underbrace{\Pr(y=B)}_{\text{constant}}\cdot\max\left\{\overbrace{\underbrace{\phi_-^B}_{\text{constant}} - \underbrace{c}_{\text{strictly increasing}}}^{>0}, 0\right\}}_{\text{strictly decreasing}}$$
$$- \underbrace{(1-\mu)}_{>0}\underbrace{(q-\pi)}_{>0}\underbrace{\left(\sigma_R^* - \sigma_N^*\right)}_{\text{constant}}.$$

Therefore, $U_V^R - U_V^N$ is strictly decreasing in c and there exists an unique $\hat{c}(\gamma)$ above which the Rational attention decreases voter welfare. QED.

[3]If σ_R^* is not constant, it is strictly increasing in c and overall expected utility difference is trivially strictly decreasing in c because of the last term.

Appendix C

CHAPTER IV

Proof of Claim 1

Proof If the low-type agent is treated as a high type for effort level e_1, her expected utility equals to:

$$EU_L^H = \underbrace{-\Big[P(i=0|s=0)(1-x_{10}) + P(i=1|s=0)(1-x_{11})\Big]}_{B_L^H} - \underbrace{ce_1^2}_{C_L^H}.$$

The cost C_L^H is trivially increasing in effort e_1. Therefore we only need to show that B_L^H is decreasing in e_1. Substituting the values of x_{10}, x_{11} and $m(e_1)$ and derivating it with respect to e_1 gives us:

$$\frac{\partial B_L^H}{\partial e_1} = \frac{e_1\theta(2-6\theta+4\theta^2)}{(-1+e_1^2(0.5-\theta)^2)^2} < 0.$$

Observe that for $\theta \in (0.5, 1)$ we always have $2 - 6\theta + 4\theta^2 < 0$ and therefore the benefit as well as overall expected utility of the low agent when treated as high is decreasing in e_1.

Proof of Proposition 24

Proof We have already discuss a few important observations in the main text needed for this proof:

1 In equilibrium, low type exerts no effort $e_0^n = 0$.

2 EU_L^H is decreasing in e_1.

3 e_1^n is the lowest effort level for which the principal believes the agent is a high type.

4 In separating equilibrium $EU_L^L(e_0^n) \geq EU_L^H(e_1^n)$, with equality if the off path beliefs satisfy the Intuitive Criterion.

First we prove that separating equilibrium cannot exist for $c < \bar{c}$. From observations above we know that in order for the low type not to have an incentive to deviate in equilibrium, we must have $EU_L^L(e_0^n = 0) = -\theta \geq EU_L^H(e_1^n)$. Moreover, $EU_L^H(e_1^n = 0) = -(1-\theta) > -\theta = EU_L^L(e_0^n = 0)$, $EU_L^H(e_1^n)$ is decreasing in e_1^n and $EU_L^L(e_0^n = 0) = -\theta$ is constant in e_1^n. Consequently, the condition for the low type stated above can never be satisfied if $EU_L^H(e_1^n = 1) > -\theta$ which implies that in order for the separating equilibrium to exist, we must have $c > \bar{c}$. Since, $EU_L^H(e_1^n)$ is strictly decreasing in e_1^n, this also proves that there is an unique value of $e_1^n \in (0, 1]$ for separating equilibrium satisfying the Intuitive Criterion.

We have already shown that these equilibrium quantities are feasible, and the low type has no incentive to deviate. As for the high type, since EU_H^H is decreasing, she never has the incentive to deviate to the effort levels where the principal correctly believes her type $e > e_1^n$. When the high type is treated as a low type, $EU_L^L(e_0^n = 0) = EU_L^H(e_1^n) < EU_H^H(e_1^n)$[1] and $EU_L^L(e_0^n = 0) = EU_H^L(e_0^n = 0) = -\theta$. This means that if the high type was considered to be a low type, she does not want to deviate to $e_0^n = 0$. We now only need to show that for $c > \bar{c}$, $EU_H^L(e)$ is a decreasing function so the high type does not want to deviate to $e \in (0, e_1^n)$.

$$EU_H^L = -\underbrace{\Big[P(i=0|s=1)(1-x_{00}) + P(i=1|s=1)(1-x_{01})\Big]}_{B_H^L} - \underbrace{ce_0^2}_{C_H^L}.$$

Unlike the case with the low type treated as high, the benefit of the high type treated as a low B_H^L is actually increasing in e_0. This is again because of the fact that the high type expects public signal to increase the policy choice. Overall,

$$\frac{\partial EU_H^L}{\partial e_0} = \frac{e_0}{(-1+e_0^2(0.5-\theta)^4}\Bigg(c\underbrace{\left(-2+4e_0^2(0.5-\theta)^2 - 2e_0^4(0.5-\theta)^4\right)}_{<0} + \underbrace{\theta(-2+6\theta-4.\theta^2)}_{>0}\Bigg)$$

Note that this value is decreasing in cost and increasing in $e_0 \in (0, 1)$. Therefore we simply check that $\frac{\partial EU_H^L}{\partial e_0} < 0$ for $c = \bar{c}$ and $e_0 = 1$, which is always the case for $\theta \in (0.5, 1)$. This proves that the high type would not want to deviate to any effort level $e \in (0, e_1^n)$. [2]

[1] Last inequality is satisfied for any positive effort level since the high type expects public signal to agree with her private signal while the low type expects the public signal to point towards lower state of the world.

[2] An interesting observation to make is that, since the derivative is increasing in $e_0 \in (0, 1)$, EU_H^L cannot have an interior maximizer for $e_0 \in [0, 1]$.

Now we only need to verify that this equilibrium satisfies the Intuitive Criterion. Since EU_L^H is decreasing in e and e_1^n makes the low type indifferent, any belief $e \in (0, e_1^n)$ is not equilibrium dominated. Therefore, the principal's beliefs that the agent is a low type when deviating to $e \in (0, e_1^n)$ satisfies the Intuitive Criterion. For $e > e_1^n$, we have equilibrium dominated for both types and therefore any belief would satisfy Intuitive Criterion on this region. This completes our proof.

Proof of Claim 2

Proof Since being considered as a low type is the worse possible case for the agent, in pooling equilibrium, the low-type agent should be weakly better off than at her ideal point when considered as low type. We have already established that the low agent when perceived as low maximizes her expected utility when $e = 0$, and this maximum utility is $EU_L^L(e = 0) = -\theta$. In pooling equilibrium, low-type agent's expected utility is

$$EU_L^P = \underbrace{-\Big[P(i = 0|s = 0)m(e) + P(i = 1|s = 0)(1 - m(e))\Big]}_{B_L^P} - ce^2.$$

If we simplify this expression, we get that $B_L^P = -0.5 + e^2(0.125 - 0.25\theta)$ which is strictly decreasing in $e \in ([0, 1]$ since $\theta > 0.5$. This result is not surprising since the decision is made just based on the public signal and low type expects public signal to be low. For higher effort level, low public signal will decrease policy choice even more. Since the cost of effort is increasing, the low type wants to exert as little effort as possible in pooling equilibrium.

EU_L^P is decreasing in e and c and we must have $EU_L^P \geq -\theta$. For $c \leq 0.125(-3+6\theta)$, this inequality is satisfied for any e. Since we are considering the best possible equilibrium for the principal, we could have pooling at the maximum level without violating the incentive compatibility of the low type when considered low. For $c > 0.125(-3 + 6\theta)$, the inequality is satisfied i.f.f. $e \leq 2\sqrt{\frac{-1+2\theta}{-1+8c+2\theta}}$[3].

Proof of Proposition 25

Proof We have already shown that this strategy is feasible and the low-type agent has no profitable deviation. We only need to make sure that high type when considered

[3]Observe that this value is always between 0 and 1 for $c > 0.125(-3 + 6\theta)$ and $\theta > 0.5$.

as low does not want to deviate from the equilibrium pooling level of effort. In previous proofs we have already established that if the high type is considered as low, her expected utility is maximized at the border $e_0 = 0$ or $e_0 = 1$. Therefore, we only need to check incentive compatibility of the high type at these two effort levels.

$EU_H^L(e_0 = 0) = EU_L^L(e_0 = 0) = -\theta$ and for the same effort level, in pooling equilibrium the high type gets strictly higher utility $EU_L^P(e) < EU_H^P(e)$. Therefore, since e^* makes the low type weakly better off than exerting no effort and being considered as low, the high type will also be better off for this equilibrium pooling level e^*.

Similar to previous proposition, it is easy to verify that for $c > 0.125(-3 + 6\theta)$, $EU_H^L(e_0)$ is decreasing in e_0, therefore we do not need to compare equilibrium expected value to $e_0 = 1$. For $c > 0.125(-3 + 6\theta)$, $EU_H^L(e_0)$, we have $e = 1$ and comparing to $e_0 = 1$ is trivial since for a fixed effort level both types are better if the principal disregards the private reports than considers her to be the low type. Therefore, high type has no profitable deviation.

Now we show that for $c < \bar{c}$ this pooling equilibrium satisfies the Intuitive Criterion. For such a cost level, we know that for any effort level e, the low type is strictly better off being considered a high type than exerting no effort and being considered as low type (or exert equilibrium pooling level). Therefore, for low type all of the effort levels are not equilibrium dominated and our beliefs satisfy the Intuitive Criterion.

Proof of Lemma 26

Proof The objective function of the principal is to maximize:

$$\max_{e\in[0,1]} \quad EU_P^r = -P(i = 1)\Big[P(\omega = 1|i = 1)(1 - x_1)^2 + P(\omega = 0|i = 1)x_1^2\Big] -$$
$$-P(i = 0)\Big[P(\omega = 1|i = 0)(1 - x_0)^2 + P(\omega = 0|i = 0)x_0^2\Big].$$

From earlier we already know that $x_0 = 1 - m(e)$ and $x_1 = m(e)$. Simplifying this objective function and substituting the values of x_0, x_1 and $m(e)$, we get that $EU_P^r(e) = -0.25 + 0.0625e^2$. This function is increasing in e for $e \in [0, 1]$ and it is trivially maximized at $e^r = 1$.

Proof of Lemma 27

Proof The Incentive Compatibility constraints for each realization of the private signal s are:

$$IC_1: \quad EU_L^L(e_0) = -P(i=0|s=0)(1-x_{00}) - P(i=1|s=0)(1-x_{01}) - ce_0^2 \geq$$
$$-P(i=0|s=0)(1-x_{10}) - P(i=1|s=0)(1-x_{11}) - ce_1^2 = EU_L^H(e_1);$$
$$IC_2: \quad EU_H^H(e_1) = -P(i=0|s=1)(1-x_{10}) - P(i=1|s=1)(1-x_{11}) - ce_1^2 \geq$$
$$-P(i=0|s=1)(1-x_{00}) - P(i=1|s=1)(1-x_{01}) - ce_0^2 = EU_H^L(e_0).$$

We have already established the following facts:

1) $EU_L^L(e_0), EU_L^H(e_1)$ and $EU_H^H(e_1)$ are decreasing in effort levels.
2) $EU_H^L(e_0)$ is decreasing in e_0 for high enough cost, and it is always maximized at the border.

If we substitute values of x_{ij} and $m(e)$ is the objective function we trivially get that the expected utility of the principal is increasing in both e_0 and e_1. Meaning, the principal who imposes different effort levels, always wants to increase the effort levels as much as possible after each private signal (if the agent is truthful).

Now we prove why the IC of the low type is binding for optimal effort levels. There are several possible cases:

- Both IC constraints are binding. If both constraints bind then we should have $EU_L^L(e_0) + EU_H^H(e_1) = EU_L^H(e_1) + EU_H^L(e_0)$. If we simplify this condition we get

$$(e_1^2 - e_0^2)\theta(1 - 3\theta + 2\theta^2) = 0.$$

  Since for separating case we have $e_1 \neq e_0$, both effort levels are positive and $\theta \in (0.5, 1])$, this condition can never be satisfied. Therefore, both IC constraints cannot bind.

- Trivially both IC constraints cannot be slack. In this case we could increase the effort level e_0 or e_1 (whichever is not at the maximum level) by ϵ small enough and it will not violate either IC constraints but will increase the value of the objective function.

- IC constraint of the high agent is binding and IC constraint of the low agent is slack. We can show that IC of the high type can only be binding for $c > -1 + 2\theta$. If the cost is lower than this value even for the highest (worst) effort level e_1, the high type strictly wants to be truthful.[4] For this high cost, $EU_H^L(e_0 = 0)$ is decreasing in e_0. Now we cannot have $e_0 = 1$ otherwise the IC constraint of the low agent would be violated (it is the worst case to be considered a low type and be forced to exert maximum effort). In this case, by increasing e_0 with ϵ small enough would not violate IC constraint of either type but increase the objective function. Therefore, in equilibrium this case is not possible either.

The only possible case left is that when the principal imposes different effort levels depending on the private report, for the optimal effort levels, the IC constraint of the low type is binding and IC of the high type is slack.

Corollary 27 When the principal imposing formal requirements induces the separation, the optimal strategy requires high agent to exert maximum effort, i.e., $e_1 = 1$.

Proof This directly follows from the previous lemma. If $e_1 < 1$, we can increase e_1 with ϵ small enough. It will not violate IC low since $EU_L^H(e_1)$ is decreasing in e_1 and will not violate IC high since the incentive compatibility constraint of the high agent is slack. However, increasing e_1 would increase the objective function giving us the contradiction.

Proof of Proposition 29

Proof We have already established that in this case, the IC constraint of the low-type agent is binding, the IC constraint of the low-type agent is slack, and $e_1 = 1$. The low-type agent's IC constraints becomes:

$$EU_L^L(e_0) = -q - ce_0^2 = B_L^H(e_1 = 1) - c = EU_L^H(e_1 = 1)$$

Observe that RHS is constant in e_0 while the LHS is strictly decreasing. Moreover, for $e_0 = 1$ we trivially have $EU_L^L(e_0 = 1) < EU_L^H(e_1 = 1)$, while $EU_L^L(e_0 = 0) >= EU_L^H(e_1 = 1) \iff c > \bar{c}$.[5] Therefore, separation is only possible for $c > \bar{c}$ and

[4] $EU_H^L(e_0)$ is maximized at the border. We trivially have $EU_H^H(e_1 = 1) > EU_H^L(e_0 = 1)$ since being considered a high type for the same effort level is strictly better. We have $EU_H^H(e_1 = 1) > EU_H^L(e_0 = 0)$ for $c < -1 + 2\theta$.

[5] Observe that this is the same threshold as before and allows the low agent not to have an incentive to deviate with $e_0 = 0, e_1 = 1$.

there is an unique value of $e_0 \in [0,1)$ that makes the low type indifferent. This proves that e_0^r is feasible and the IC constraint of the low type is satisfied. We also check that for $c > \bar{c}$ we also have $EU_H^H(e_1 = 1) > EU_H^L(e_0 = e_0^r)$. The effort levels described in the proposition are feasible, satisfy both incentive constraints, and are optimal based on the previous lemma and its corollary.

Proof of Proposition 30

Proof We have already shown the optimal strategies for separating and pooling cases. We need to show that the principal has no profitable deviation (from pooling to separating and vice versa). The principal's expected utility in pooling equilibrium does not depend on cost. Substituting the optimal maximum effort level, the principal in the optimal pooling case gets $EU^p(e^r = 1) = -0.1875$.

The principal can induce separation i.f.f. $c > \bar{c}$. Therefore, the equilibrium below this cost level is simply requiring maximum effort level after each private report. For $c > \bar{c}$, the principal's expected utility playing the optimal strategy with separation is $EU(e_0^r(c), 1)$. Even though the principal does not bare the cost of effort directly, this expected utility depends on c through the optimal effort after the low report $e_0^r(c)$.

We know that $e_0^r(c = \bar{c}) = 0$ and $e_0^r(c)$ is increasing in c. Since the principal always prefers more effort, $EU(e_0^r(c), 1)$ is also increasing in c. Moreover,when $c \to \infty \implies e_0^r(c) \to 1 \implies EU(e_0^r(c), 1) \to EU(e_0 = 1, e_1 = 1) > EU^p(e^r = 1)$. While $EU(e_0^r(\bar{c}), 1) = EU(0, 1) \leq -0.1875 \iff \theta \leq 0.698539 \equiv \bar{\theta}$.

Therefore, we get that for $\theta > 0.698539$, the optimal strategy in separation is always strictly better than the optimal strategy in pooling for the principal (whenever the separation is possible $c > \bar{c}$). For $\theta \leq 0.698539$, there is an unique $\tilde{c} > \bar{c}$ such that $EU(e_0^r(c), 1) > EU^p(e^r = 1) \iff c > \tilde{c}$.