The London School of Economics and Political Science

Essays in information economics

Clement Minaudier

A thesis submitted to the Department of Economics of the London School of Economics and Political Science for the degree of Doctor in Philosophy

London, April 2019

Declaration

I certify that the thesis I have presented for examination for the PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

I declare that my thesis consists of approximately 39,000 words.

Acknowledgments

I am deeply grateful to my supervisor Gilat Levy for her invaluable guidance and constant support, and to my advisors Ricardo Alonso and Stephane Wolton for precious advice and encouragement throughout my PhD.

I would also like to thank the microeconomic theory group: Andrew Ellis, Erik Eyster, Mathew Levy, Francesco Nava, Ronny Razin, and Balazs Szentes for encouraging me to pursue the topics I studied in this thesis and for their advice.

I am grateful to the Political Science and Political Economy group for welcoming me to present my research in their seminars and for providing valuable advice. This includes in particular Ali Cirone, Torun Dewan, Marco Giani, Rafael Hortalla-Valve, Valentino Larcinese, Janne Tukiainen, Thomas Leeper, and Tinghua Yu.

I am equally indebted to my fellow research students in both the economics and government departments not only for their academic advice but also for their friendship and support. In particular, I would like to thank Shan Aman-Rana, Thomas Brzustowski, Weihan Ding, Simon Franklin, Alkis Georgiadis-Harris, Selina Hofstetter, Federica Izzo, Krittanai Laohakunakorn, Claudia Robles-Garcia, Marcus Roel, Nelson Ruiz, Francesco Sannino, Greg Sheen, Xuezhu Shi, Manuel Staab, Kostas Tokis, Arduino Tomasi, Heidi Thysen, and Lisa Windsteiger.

I also thank researchers outside the LSE with whom I have had the chance to discuss some of these projects, including Andy Guess, Yingni Guo, Niall Hughes, Anton Kolotilin, and Chris Tyson.

Finally, I would like to thank Mark Wilbor and Deborah Adams, for their invaluable help throughout my time as a PhD student.

My family and friends have been a source of constant support, and I am grateful to all of them beyond words. This thesis would not have been possible without the support and encouragement of my fiancée Vanessa.

Financial support from the Economic and Social Research Council (grant number ES/J500070/1) is gratefully acknowledged.

Abstract

These essays examine how economic agents strategically choose to produce, manipulate, or disclose information, when that information can influence the behaviour of others. By theoretically modelling these choices, it seeks to contribute to debates about the optimal design of policies such as transparency rules, the regulation of lobbying, or the concentration of ownership among information providers such as media groups. The models developed in these essays also provide a framework to interpret and evaluate empirical assessments of how information influences behaviour.

The first chapter looks at how interest groups choose to generate information to influence policies. It innovates on the literature by explicitly modelling the choice of policy makers to obtain their own confidential internal information ahead of interactions with these groups. This approach reveals unintended consequences of transparency policies and the subtle role that institutions such as congressional research agencies can have on the quality of policy making.

The second chapter studies how agents choose to produce new information, for instance by running experiments, in the presence of competing information providers. In particular, it examines whether these agents produce more information when they compete than when they collude. The existing literature has established that when these agents possess no existing information, competition always increases the amount of new information produced. I show that when agents do possess prior information, this conclusion does not necessarily hold.

The third chapter analyses how policy choices are affected when voters have a limited capacity to correctly interpret information about policy performance. In a situation where policy performance provides information about the competence of policy makers, and where voters decide whether to re-elect incumbents based on that information, voters may benefit from these cognitive limitations as they can induce policy makers to choose better policies.

Contents

List of Figures

Chapter 1

The value of confidential policy information: persuasion, transparency, and influence

1.1 Introduction

Transparent policy making is often considered a defining feature of democracy.[1] When the information available to policy makers is easily accessible, the public can scrutinise policy decisions and hold elected representatives accountable.

While governments have started disclosing the identity of external sources of information and the interests they represent, through bills such as the Lobbying Disclosure Act in 1995 in the US or the Transparency of Lobbying Act in the UK in 2014, they continue to defend forcefully the confidentiality of internal sources. For example, the Congressional Research Service (CRS) in the US has strictly restricted access to its research for the last 30 years.[2] One of the arguments advanced by the CRS to defend the confidentiality of its reports is the risk of influence by outsiders: "Widespread public dissemination will almost certainly increase partisan and special interest pressure [...]. Such pressure from the public [...] could subtly affect the way CRS authors write their reports. Congress may ultimately benefit less from the information in CRS Reports."[3]

This argument suggests that there can be a cost associated with the transparency of internal information. This cost is closely tied to the co-existence of internal and external sources of information, and to the influence that special interest groups can exert on the policy process. While internal information is obtained through formal institutions such as parliamentary research services, government agencies, or committees (Howlett 2015), external information is generally provided by lobbyists on behalf of special interest groups (Esterling 2004). In 2016, lobbying expenditures in the US amounted to \$3.15 billion. These resources, used to transmit information to policy makers, are much larger than campaign contributions – which amounted to less than \$500 million for the whole 2015/2016 electoral cycle – and therefore represent an important channel of influence.[4]

In this paper, I evaluate the effect of keeping internal information confidential on the provision of external information. I extend theories of informational lobbying – the influence of interest groups through the provision of information, rather than through monetary contributions – by explicitly considering policy makers' control over their inter-

[1] As an example, Angel Gurría, OECD Secretary–General stated in a recent speech that "The OECD has been at the forefront of efforts to promote and protect the free flow of information. We believe this to be a fundamental human right." (Source `OECD.org`).

[2] This was formally expressed in 1980 when, in response to a subpoena by the Federal Trade Commission on behalf of oil companies to access all CRS research related to the oil industry, the Senate issued a resolution stating that "The communications of the Congressional Research Service to the members and committees of the Congress are under the custody and control of the Congress." (S. Res. 396, 96th Cong., 2d Sess. 1980, cited in `https://fas.org/sgp/crs/crs041807.pdf`).

[3] See `https://fas.org/sgp/crs/considerations.pdf`.

[4] Campaign contributions from PACs only, excluding individual donations. Source: Center for Responsive Politics.

nal information. This approach reveals a novel channel by which confidentiality can be beneficial: by keeping their own information confidential, policy makers can induce special interest groups to provide more evidence. The value of confidentiality to policy makers is not driven by reputational concerns or bargaining considerations, and can therefore be socially beneficial. Characterising policy makers' strategic choices of internal information allows me to derive testable predictions on the value of confidentiality, and to show that empirical assessments of influence should account for the role of government expertise.

The provision of information by special interest groups depends on policy makers' choices of internal information and on whether that information is confidential. When interest groups can observe the information already available to policy makers, they can produce evidence that is just sufficiently accurate to tilt the policy decision in their favour. When the information available to the government is not publicly available, interest groups form beliefs about the information policy makers are most likely to have. These beliefs determine whether interest groups want to offer more or less information: if they believe that policy makers are likely to be sceptical about their preferred policy, then they need to offer more evidence. Therefore, policy makers should shape their preliminary investigations to let lobbyists believe that they are sufficiently sceptical and that more evidence is needed.

I formalise this intuition to address the following questions. First, when is confidentiality valuable to policy makers, and therefore most likely to be used by governments? Second, how does the government's control over internal investigations affect the influence that special interest groups exert on policy making?

To answer these questions, I consider a model with a single policy maker and a lobbyist. The policy maker has to decide whether to enact a new policy that is supported by the lobbyist, but faces uncertainty. In the first stage, the policy maker chooses the precision of a signal about an unknown state of nature that she receives confidentially. She would like to choose the welfare-maximising policy given the state, but her limited expertise constrains the precision of her signal. The lobbyist also acquires some independent signals and commits to revealing them to the policy maker. This allows him to engage in Bayesian persuasion: his expertise is not limited and he can perfectly adjust the precision of his information to persuade the policy maker to choose his preferred policy.

I show that confidentiality is valuable as it forces the lobbyist to choose an investigation that reveals more evidence than necessary to persuade the policy maker. By keeping her own signal realisations confidential, the policy maker strategically creates a situation of asymmetric information which allows her to extract informational rent. This occurs even when her preliminary information would have no effect on her policy choice, in the absence of lobbying.

This result can explain why governments sometimes insist on the confidentiality of their information, while lobbyists criticise the lack of access to that information. For instance, when the UK government refused to publish studies evaluating the impact of Brexit between 2016 and 2017, one of the groups it received evidence from, the Food and Drink Association, stated that "[The] Government has a duty to share this analysis with the sector so businesses can prepare".[5]

A puzzling aspect of this confidentiality is that this information does not always reveal some government wrongdoing, or weakens the government's negotiating position. For instance, the research produced by the CRS is available to the whole congress, and therefore cannot affect the bargaining power of a legislator against other members of congress. Similarly, while some members of the UK government worried that revealing information on the impact of Brexit might have weakened the government's bargaining position with the European Union, it also seems unlikely that EU negotiators did not already possess similar information.[6] Strengthening the government's negotiating position therefore does not seem to be the only reason for the confidentiality of these studies. This model suggests an explanation for this puzzle by showing that confidentiality can be valuable even in the absence of reputational concerns.

The value of confidentiality has limits, however. I show that the policy maker may need to distort her own information, in order to induce the lobbyist to provide more evidence than he would like to. These distortions involve reducing the precision of certain conclusions of the investigation, and hence reduce its overall informativeness. There is thus a trade-off between obtaining information internally and extracting it from external sources.

A second result of this analysis is therefore that the value of confidentiality is non-monotonic in the policy maker's expertise and her ideological alignment with lobbyists. When government expertise is low, the value of confidentiality increases in expertise,

[5] In a study of interest groups access to legislators in Estonia conducted by the Praxis Center for Policy Studies, over 50% of them emphasised that information from the government was not easily accessible (Jemmer 2014). There is also evidence that lobbyists adapt their strategies to the transparency of the policy process. Matthews Luxon (2012) shows that lobby groups react to lower government transparency by making their lobbying tactics more specialised. As a result, some groups specialise in offering evidence to support policy proposals when transparency is lower, while this specialisation does not occur when transparency is high.

[6] Indeed, this point was raised by a member of the opposition requesting the studies to be published when he asked in a Commons debate: "Do they [the government] honestly believe that the EU has not carried out its own assessments of what Brexit will mean for those 58 areas?" (HC Deb (01 Nov 2017)).

as more expertise allows the policy maker to extract more evidence from the lobbyist. However, when expertise is high, that value eventually declines as the additional gains from inducing the lobbyist to provide more information are relatively less important, compared to the costs of distorting internal information. The policy maker's predisposition to accept the lobbyist's preferred policy, measured by their ideological alignment, also affects the amount of evidence produced and therefore the value of confidentiality. I show that confidentiality is most valuable when the ideological alignment between the policy maker and the lobbyist is neither too low nor too high.

Finally, modeling the policy maker's choice of internal investigation reveals that the influence of the lobbyist on policy can sometimes increase in the government's expertise. This arises because more expertise can make the policy maker more likely to choose the lobbyist's least-preferred policy and therefore makes the lobbyist's presence even more important to overturn that choice. This occurs despite the fact that expertise improves policy choice and makes the policy maker better-off. Similarly, an increase in alignment can correspond to a decrease in influence – even though more alignment makes persuasion easier for the lobbyist – because it can make the policy maker more likely to choose the lobbyist's preferred policy in the absence of lobbying.

These results have both positive implications for evaluating the influence of interest groups, and normative consequences for the optimal design of institutions.

Consider the fact that the resources spent on lobbying in the US have significantly increased in real terms over the last 15 years, while the budget of the Congressional Research Service has remained relatively constant or even declined, as shown in figure 1.1. This observation could suggest that information generation is increasingly being outsourced to external groups whose influence has therefore increased over time. This paper proposes an alternative explanation: even when their capacity to gather information internally is reduced, policy makers can force lobbyists to provide additional information, and therefore expend more resources, so these changes in fact reflect a loss of influence. Which of these two explanations is correct depends on how these sources of information interact. More generally, the model reveals that careful consideration of the counterfactual policy choice in the absence of lobbying is required to correctly interpret empirical evidence on influence. Since the choice of internal investigation differs in the absence of lobbying, the default policy choice can also differ, and influence should not be simply measured as the probability of a policy change following lobbying efforts. The results also show that research designs that fail to control for government expertise may produce biased estimates of the returns to lobbying, as expertise can be correlated with both lobbying resources and the probability of a policy change.

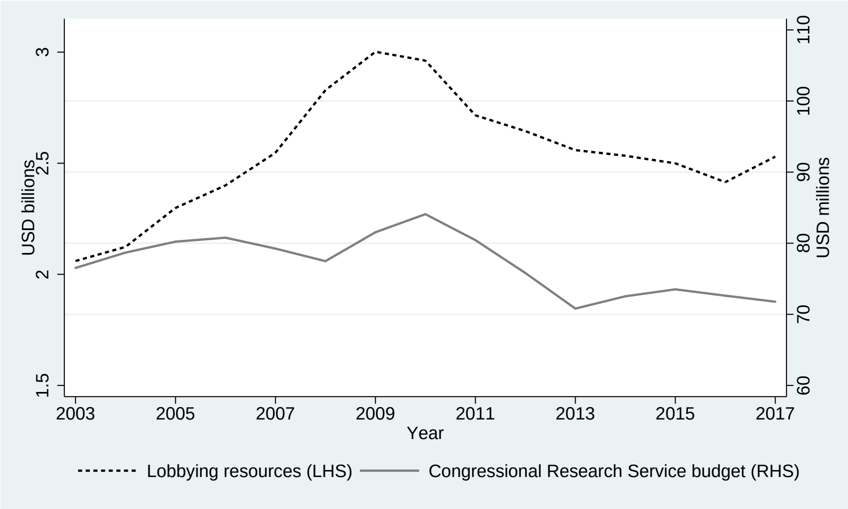


Figure 1.1: Lobbying resources (LHS) and Congressional Research Service budget (RHS), inflation-adjusted. Sources: CRS annual report and Center for Responsive Politics.

Secondly, the results explain why policy makers may prefer their information to be confidential in certain policy areas but publicly available in others, and how the choice of confidentiality can vary over time as the policy agenda evolves or new, more technocratic, policy makers get elected. These strategic preferences for confidentiality can be positively correlated with the quality of policy making even in the absence of a causal relationship between the two variables. Islam (2006) reports that higher levels of transparency are associated with better governance and concludes that "there is a strong positive relationship between transparency and governance, with the likely effect running from the former to the latter". Focusing on the effect of government transparency on informational lobbying suggests that there can be other factors, such as government expertise, moving both of these variables without the existence of a causal link between them. In fact, the model shows that shining too much light on the policy process can reduce the quality of policy making, and therefore have adverse consequences beyond those already identified in the literature (see e.g. Prat 2005).

Related literature

This paper relates to two strands of literature: models of informational lobbying and studies of transparency in political institutions. It shows that these two questions are linked: transparency determines how information should be allowed to flow between policy makers and lobbyists which, in turn, affects the type of information lobbyists choose to provide.

A large literature has looked at how information is transmitted to legislators by lobbyists (e.g. Potters & van Winden (1992), Austen-Smith & Wright (1992), Rasmusen (1993), Austen-Smith (1993), Lagerlof (1997)). The most closely related papers within that literature study how informational lobbying is affected by information already held by policy makers. Felgenhauer (2013) shows that expert politicians are not always better at making decisions than non-experts in the presence of lobbyists. In his model, the expertise of the politician cannot affect the information provided by a single lobby and only has an effect when two lobbies compete. By allowing the information to be concealed, I show that even a single lobby can be induced to provide more information as the politician's expertise increases. Cotton & Dellis (2016) show that informational lobbying can be detrimental if more information provided by lobbyists shifts the focus of a policy maker towards less important issues and thus reduces the information she collects. This substitution across the two sources of information relies on the existence of multiple policies and the limited capacity of the policy maker to act on these different policies. Substitution arises in my model even with one policy dimension because information can be confidential, so that the policy maker's choice of information affects the beliefs of lobbyists and the amount of evidence they want to provide. Finally, in Ellis & Groll (2017), the trade-off between acquiring costly information in-house or relying on that provided by lobbyists comes from the difference in resource constraints of these two sources. Information is costless in my model and the interaction between the two types of information relies on whether that information is made public or not.[7] Another closely related paper, Cotton & Li (2018), studies the effect of internal information on monetary lobbying. They show that because a better informed politician might be harder to sway through contributions, politicians might prefer to remain uninformed or to reduce the informativeness of the signals they obtain. While they share some of the implications of this paper, they focus on the effect of internal information on monetary contributions rather than on information provision. Since influence can take both forms, this paper is complementary to theirs. With informational lobbying, additional information from the politician can be detrimental even when the politician wants to choose the socially optimal policy rather than to maximise contributions.

[7] Other papers also study how political institutions affect the influence of informational lobbying. Bennedsen & Feldmann (2002a) look at the effect of the vote of confidence procedure, Bennedsen & Feldmann (2002b) at party cohesion, Dellis & Oak (2018) at the legislature's subpoena power, while Dahm & Porteiro (2008a) and Wolton (2018) look at the interaction between informational lobbying and other forms of pressure. Finally, Schnakenberg (2017) shows how the presence of multiple legislators who can communicate affects informational lobbying.

In the transparency literature, Felgenhauer (2010) and Gailmard & Patty (2018) study the effect of making policy makers' information public.[8] Felgenhauer (2010) finds that confidentiality can be beneficial to the public as it can induce lobbyists to refrain from providing monetary contributions and therefore result in better policy decisions. In his paper, more internal information results in both a better decision and less influence from lobbyists and is therefore unambiguously valuable. By looking at informational instead of monetary lobbying, I show that more precise information can be detrimental and can make confidentiality less valuable.[9] Gailmard & Patty (2018) find that transparency of the policy maker's information can reduce the amount of information transmitted from a bureaucrat and show that the policy maker prefers opacity over transparency if preferences are sufficiently different. By focusing on delegation in bureaucracies, rather than interest group influence, they introduce a trade-off between authority and information aggregation which leads to a focus on the choice of delegation rule rather than distortions in the policy maker's information.

From a technical perspective, this paper is related to the literature on Bayesian persuasion. Kamenica & Gentzkow (2011) identify the optimal signal that a sender can design to persuade a single receiver, assuming that both sender and receiver are symmetrically informed about the state of the world. The results developed in their paper are used throughout this paper. A number of recent papers look at how these results change when the receiver is privately informed.[10] The most closely related studies in this literature are Guo & Shmaya (2018) and Kolotilin (2018). In both papers, a sender chooses an optimal information structure to persuade a privately informed receiver. Guo & Shmaya (2018) provide a general solution for the sender's persuasion strategy when the receiver's information is correlated with the state. Kolotilin (2018) restricts attention to the case where payoffs are linear in types and the receiver's type is uncorrelated with the state and shows that, under certain conditions, the receiver's payoff can be decreasing in the precision of her private information. My results show that, if the receiver were to choose her private information, she would indeed prefer to keep it private as long as its precision is limited.

[8] A large literature has shown how transparency of an agent's *actions* can be damaging in a number of institutions, including decision making in committees of experts (Levy (2007), Meade & Stasavage (2008), Seidmann (2011), Swank & Visser (2013), Hansen et al. (2017), Fehrler & Hughes (2018), Gradwohl & Feddersen (2018)) or by a single expert (Fox & Van Weelden (2012)), political accountability (Fox (2007), Stasavage (2007), Malesky et al. (2012), Carey (2013) Benesch et al. (2018)), international negotiations (Stasavage (2004), Naurin (2007)), or more general principal-agent relationships (Prat (2005)).

[9] Another key difference is that monetary contributions have no social value so the trade-off faced by the politician in Felgenhauer (2010) is between accepting the lobbyist's contribution and choosing the socially valuable policy.

[10] See for instance, Li & Shi (2017) in a bilateral trade setup, Basu (2017) in a dynamic setting, and Au (2015) and Kolotilin et al. (2017) for the case where the receiver's private information is about her payoffs rather than about the state of the world.

The rest of the paper is organised as follows. Section 1.2 introduces the model. Section 1.3 provides the main results: it shows how the policy maker gains from confidentiality and characterises both the lobbyist's choice of persuasion strategy and the policy maker's choice of preliminary investigation. Section 1.4 shows that confidentiality is only valuable when expertise or ideological alignment are not too high, and that interest group influence varies non-monotonically in these two variables. Section 1.5 discusses the implications of these results and section 1.6 concludes. All proofs are presented in appendix.

1.2 Model

The model has two players: a policy maker and a lobbyist, and three stages. In the first stage, the policy maker can acquire some information about a binary state of the world $\omega \in \{0, 1\}$. In the second stage, the lobbyist acquires some additional information about ω to present to the policy maker. All players share a prior $\mu_0 := \mathbb{P}(\omega = 1)$. Throughout the paper a belief refers to the probability that the state is $\omega = 1$, unless otherwise specified.

In the final stage, the policy maker chooses a policy $x \in \{0, 1\}$ whose consequences are uncertain. The policy maker wants the policy to match the state. Her preferences are represented by the payoff function

$$u(x, \omega) = \begin{cases} 1 & \text{if } x = \omega, \\ 0 & \text{if } x \neq \omega \end{cases}$$

By contrast, the lobbyist cares about the final action of the policy maker independently of the state, and wants her to choose policy $x = 1$. His preferences are represented by the payoff function

$$v(x) = \begin{cases} 1 & \text{if } x = 1, \\ 0 & \text{if } x = 0 \end{cases}$$

Because of the uncertainty, the policy maker would like to obtain more information about the state. This information can come from two sources. In the first stage, she can launch a *preliminary investigation*. A *preliminary investigation* consists of a pair of conditional probability distributions over binary signal realisations $r \in \{r_0, r_1\}$, for each value of the state:[11]

$$p = \{p(r|\omega = 1), p(r|\omega = 0)\}$$

The policy maker updates her beliefs according to Bayes rule upon observing $r \in \{r_0, r_1\}$, to

$$\mu^r = \mathbb{P}(\omega = 1|r) = \frac{p(r|\omega = 1)\mu_0}{p(r|\omega = 1)\mu_0 + p(r|\omega = 0)(1 - \mu_0)}$$

The information provided by a preliminary investigation is therefore captured by the posterior beliefs it can induce: μ^{r_0} or μ^{r_1}.

In the second stage, the policy maker obtains additional information from the lobbyist. The lobbyist produces evidence in the same way as the policy maker does, by choosing an investigation which produces one of two signals s_0 and s_1. Since the lobbyist uses this investigation to attempt to persuade the policy maker, I refer to this choice of investigation as the lobbyist's *persuasion strategy* and denote it π. A *persuasion strategy* consists of a pair of probability distributions over realisations $s \in \{s_0, s_1\}$ conditional on ω:

$$\pi = \{\pi(s|\omega = 1), \pi(s|\omega = 0)\}$$

Before s is realised, the lobbyist can credibly commit to revealing it to the policy maker. The lobbyist therefore cannot conceal his evidence or lie about it. This commitment assumption is a standard feature of the Bayesian persuasion literature (see e.g. Kamenica & Gentzkow 2011) and has been used to model informational lobbying (e.g. Austen-Smith 1998, Cotton & Dellis 2016). Given the focus of this paper on the policy maker's ability to mitigate influence, allowing the lobbyist to have full commitment provides the most demanding benchmark: it captures a situation in which the lobbyist has a significant advantage and where influence is hardest to mitigate.[12]

The lobbyist's strategic choice is therefore over which of many possible signal structures to choose to affect the policy maker's posterior beliefs. I denote the posterior belief of the

[11] In the appendix, I discuss the possibility that the policy maker has access to a more complex investigation generating more than two signals and show that the main insights continue to hold.

[12] There are a number of situations in which this assumption is satisfied. For instance, special interest groups may fund and help design scientific studies. Once the results of these studies are released in peer-reviewed publications, special interest groups can no longer control their disclosure (for examples, see White & Bero 2010, Kearns et al. 2016, Nestle 2016). Pharmaceutical companies have also funded patient advocacy groups to send patients to testify in Congress (Kopp et al. 2018). The companies can influence what patients are likely to reveal, but do not have control over the final testimony, so this type of influence strategy is akin to running an uncertain experiment and committing to disclosing the results.

policy maker following realisations r and s by $\mu_s^r := \mathbb{P}(\omega = 1|s, r)$.

Acquiring information is costless for both players. However, a key feature of the model is that the policy maker's expertise, i.e. her capacity to produce information, is limited. Formally, expertise is captured by a bound $B \in [1, +\infty)$ on the likelihood ratios of the signals r. So that, for every p,

$$\frac{1}{B} \leq \frac{p(r|\omega)}{p(r|\omega')} \leq B$$

This bound implies that the policy maker cannot learn the state of the world perfectly: the posterior beliefs μ^r must belong to an interval $\left[\underline{\mu}, \bar{\mu}\right] \subset [0, 1]$. The lowest and highest posterior beliefs that she can induce are: $\underline{\mu} = \frac{\mu_0}{\mu_0 + (1-\mu_0)B}$ and $\bar{\mu} = \frac{B\mu_0}{B\mu_0 + (1-\mu_0)}$. The parameter B captures the difference in expertise between the policy maker and the lobbyist. The lobbyist's advantage stems from facing no expertise bound as, in effect, $B = +\infty$ for him.

I refer to the preliminary investigation p such that $\frac{p(r_0|\omega=0)}{p(r_0|\omega=1)} = \frac{p(r_1|\omega=1)}{p(r_1|\omega=0)} = B$ as the *the most informative preliminary investigation* available and denote it $\bar{p}$. This investigation induces interim beliefs $\mu^{r_0} = \underline{\mu}$ and $\mu^{r_1} = \bar{\mu}$. When the policy maker chooses that preliminary investigation, I say that she *makes full use of her expertise.*

Under confidentiality, the policy maker's choice of preliminary investigation p is publicly observed by the lobbyist, but not its outcome r.[13] The choice of persuasion strategy π can therefore be conditioned on p but not on r.

The timing is as follows:

1. The policy maker publicly chooses a preliminary investigation p.
2. $r \in \{r_0, r_1\}$ is realised but only observed by the policy maker.
3. The lobbyist chooses a persuasion strategy π after observing p.
4. $s \in \{s_0, s_1\}$ is publicly realised.
5. The policy maker updates her beliefs and chooses $x \in \{0, 1\}$.

Under transparency, the timing is the same, but the lobbyist can observe the realised r and therefore condition π on both p and r.

[13]This corresponds to examples mentioned in the introduction in which the type of information obtained by the government was known, but the results were kept confidential. More generally, policy makers can run pilot projects, or commission reports from the civil service in visible ways without publicising the results of these investigations. In the appendix, I show that the policy maker does not need to commit to keeping that information confidential: if she had the possibility to disclose it, she would never choose to do so in equilibrium.

The equilibrium concept is weak perfect Bayesian equilibrium: the players' strategies are sequentially rational given their beliefs, and beliefs are updated according to Bayes rule whenever possible.[14] I focus on pure strategy equilibria within this class.[15]

Policy choice

I begin by solving for the policy maker's choice of policy x in the final stage of the game, given some generic belief μ. The policy maker chooses policy $x = 0$ (respectively, $x = 1$) if she is sufficiently confident that the state is $\omega = 0$ (respectively, $\omega = 1$). I assume that the policy maker selects $x = 1$ when indifferent. The policy choice can be expressed as a function of some generic posterior belief μ:

$$x(\mu) = \begin{cases} 0 & \text{if } \mu < \frac{1}{2}, \\ 1 & \text{if } \mu \geq \frac{1}{2} \end{cases}$$

Given this strategy $x(\mu)$, we can express the policy maker and lobbyist's expected utilities as functions of μ. Let $U(\mu) = \mu u(x(\mu), 1) + (1 - \mu)u(x(\mu), 0)$ and $V(\mu) = v(x(\mu))$. For the policy maker,

$$U(\mu) = \begin{cases} 1 - \mu & \text{if } \mu < \frac{1}{2}, \\ \mu & \text{if } \mu \geq \frac{1}{2} \end{cases}$$

While for the lobbyist,

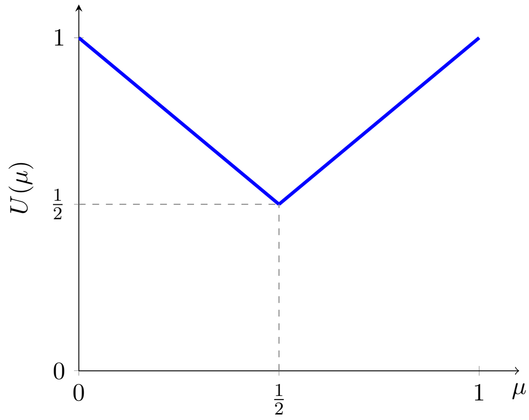
$$V(\mu) = \begin{cases} 0 & \text{if } \mu < \frac{1}{2}, \\ 1 & \text{if } \mu \geq \frac{1}{2} \end{cases}$$

These expected utilities are illustrated in figure 1.2.

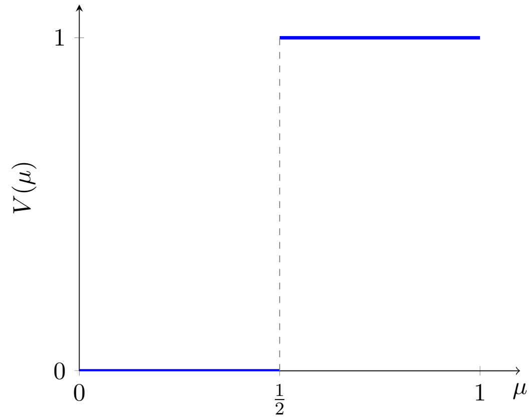
I focus on the more interesting case where the policy maker needs to be persuaded to take action $x = 1$, that is, I assume that $\mu_0 < \frac{1}{2}$. The closer the prior belief μ_0 is to $\frac{1}{2}$, the more sympathetic the policy maker is to the lobbyist's proposal (while remaining unfavourable to that proposal). I therefore refer to an increase in μ_0 towards $\frac{1}{2}$ as an

[14]Since players can never signal any private information through their choice of action before the other player's move, there is no need to refine beliefs following off-equilibrium actions.

[15]There cannot exist a mixed strategy equilibrium in which the policy maker mixes across policy choices in the final stage. If there was, then the lobbyist could always deviate to break the policy maker's indifference and increase the probability that she chooses his preferred policy. In addition, mixing across investigations or persuasion strategies simply leads to another distribution over posterior beliefs, which could be replicated with a different investigation or persuasion strategy.



(a) Policy maker's expected utility



(b) Lobbyist's expected utility

Figure 1.2: Policy maker and lobbyist's expected utilities

increase in the *ideological alignment* between the policy maker and the lobbyist.[16]

For a given p and a given π, the policy maker's ex-ante expected utility is

$$\mathbb{E}\left[U(\mu)|(p,\pi)\right] = \sum_{r\in\{r_0,r_1\}} \mathbb{P}_p(r) \sum_{s\in\{s_0,s_1\}} \mathbb{P}_\pi(s|r) U\left(\mu_s^r\right)$$

$\mathbb{P}_p(r)$ is the probability of observing realisation r from the policy maker's investigation p: $\mathbb{P}_p(r) = \mu_0 p(r|1) + (1-\mu_0)p(r|0)$. Similarly, $\mathbb{P}_\pi(s|r)$ is the probability of observing realisation s from the lobbyist's persuasion strategy π, conditional on having observed signal r that is: $\mathbb{P}_\pi(s|r) = \mu^r\pi(s|1) + (1-\mu^r)\pi(s|0)$.

1.3 The role of confidential policy information

In this section, I explain the core mechanism that allows the policy maker to extract information from the lobbyist by using her preliminary investigation. I first describe the players' strategies when the policy maker's information is not confidential. I then show how the lobbyist adapts his strategy to confidentiality and how the policy maker maximises the benefits from her preliminary investigation.

1.3.1 Transparency

Suppose first that the policy maker's information is not confidential, so her belief following a realisation r from her preliminary investigation is the same as that of the lobbyist: μ^r.

[16] A policy maker who is more ideologically aligned with a lobbyist is not necessarily a supporter of the lobbyist's preferred policy. This definition of ideological alignment corresponds to the idea of *proximity* between legislators and lobbyists and is in line with definitions used in the empirical literature. For example Igan & Mishra (2014) control for legislator–congress fixed effects to account for "potential changes in a legislator's general propensity to switch stances across time".

Lobbyist's persuasion strategy

The lobbyist's strategy, π, induces a lottery over the policy maker's posterior beliefs: with probability $\mathbb{P}_\pi(s_0|r)$, the policy maker will observe signal realisation s_0 and update her belief to $\mu^r_{s_0}$, while with probability $\mathbb{P}_\pi(s_1|r)$ she will update her belief to $\mu^r_{s_1}$. The posterior beliefs satisfy: $\mu^r_{s_0} \leq \mu^r \leq \mu^r_{s_1}$.

The lobbyist only gains when the policy maker chooses $x = 1$, which requires her belief to be above $\frac{1}{2}$. If the belief of the policy maker is such that she already chooses the lobbyist's preferred policy ($\mu^r \geq \frac{1}{2}$), the lobbyist does not need to provide any evidence.

If $\mu^r < \frac{1}{2}$, a posterior belief above $\frac{1}{2}$ can only occur following realisation s_1. The lobbyist's problem is therefore to maximise the probability that s_1 occurs, while ensuring that the policy maker chooses policy $x = 1$ after observing that realisation. As shown in Kamenica & Gentzkow (2011), this is achieved by choosing a persuasion strategy π such that the policy maker is just sufficiently persuaded following favourable evidence ($\mu^r_{s_1} = \frac{1}{2}$), and such that unfavourable evidence is as precise as possible ($\mu^r_{s_0} = 0$). Any additional favourable evidence (i.e. inducing $\mu^r_{s_1} > \frac{1}{2}$) would be wasted, as the policy maker already chooses the lobbyist's preferred policy. In addition, any other level of unfavourable evidence (i.e. $\mu^r_{s_0} > 0$) would not change the policy maker's decision and the lobbyist's payoff, but would make s_0 more likely, which reduces the lobbyist's expected utility.

Lemma 1. *Under transparency, if the policy maker is not already persuaded ($\mu^r < \frac{1}{2}$), the lobbyist's equilibrium persuasion strategy, denoted π_r, induces the beliefs $\mu^r_{s_1} = \frac{1}{2}$ and $\mu^r_{s_0} = 0$, and therefore satisfies*

$$(\pi_r(s_1|\omega = 1), \pi_r(s_1|\omega = 0)) = \left(1, \frac{\mu^r}{1 - \mu^r}\right)$$

Note that as μ^r increases, persuasion becomes easier ($\mathbb{P}_{\pi_r}(s_1|r) = \mu^r \pi_r(s_1|1) + (1 - \mu^r)\pi_r(s_1|0) = 2\mu^r$ increases), and the lobbyist needs to provide less information (the 'noise' from his persuasion strategy, $\pi_r(s_1|\omega = 0) = \frac{\mu^r}{1-\mu^r}$, increases). This relationship between the policy maker's belief and the lobbyist's strategy will also determine the equilibrium strategy when the policy maker has confidential information.

Policy maker's preliminary investigation

Given this persuasion strategy, the policy maker's expected utility as a function of her interim beliefs is

$$U^P(\mu^{r_0}, \mu^{r_1}) = \begin{cases} \sum_{r \in \{r_0, r_1\}} \mathbb{P}(r) \left[\mathbb{P}_{\pi_r}(s_0|r) \cdot 1 + \mathbb{P}_{\pi_r}(s_1|r) \cdot \left(\frac{1}{2}\right) \right] & \text{if } \mu^{r_1} < \frac{1}{2}, \\ \mathbb{P}(r_0) \left[\mathbb{P}_{\pi_{r_0}}(s_0|r_0) \cdot 1 + \mathbb{P}_{\pi_{r_0}}(s_1|r_0) \cdot \left(\frac{1}{2}\right) \right] + \mathbb{P}(r_1)\mu^{r_1} & \text{if } \mu^{r_1} \geq \frac{1}{2} \end{cases} \tag{1.1}$$

Where the probability of a realisation r, is determined by the pair of interim beliefs (μ^{r_0}, μ^{r_1}) and the Bayes plausibility constraint that they need to satisfy: $\mathbb{P}(r_0)\mu^{r_0} + \mathbb{P}(r_1)\mu^{r_1} = \mu_0$.

The policy maker chooses her preliminary investigation to maximise the total information that she receives. If her expertise is limited and the interim beliefs μ^r she can generate are always below $\frac{1}{2}$, the policy maker cannot gain from her preliminary information because the lobbyist will optimally adjust his persuasion strategy to the policy maker's belief. The policy maker is therefore indifferent between any preliminary investigation.

If her expertise is sufficiently high and her own investigation can persuade her ($\bar{\mu} > \frac{1}{2}$), the lobbyist provides no valuable information following r_1 if $\mu^{r_1} > \frac{1}{2}$. The policy maker faces a sharp trade-off: if she chooses the most informative investigation herself, and obtains some belief $\bar{\mu} > \frac{1}{2}$, the lobbyist stops providing information. This trade-off is resolved in favour of obtaining more preliminary information: by choosing the most informative investigation (and inducing $\bar{\mu} > \frac{1}{2}$), she can become more confident in her policy decision ($x = 1$) than she would ever be if she were to restrict her information ($\mu^{r_1} < \frac{1}{2}$) and rely on the lobbyist's information.

In both cases, the policy maker does not gain from the lobbyist's information, and it is therefore optimal for her to obtain as much preliminary information as possible.[17]

Proposition 1. *Under transparency, the most informative investigation ($\bar{p}$) is an equilibrium strategy for the policy maker.*

1.3.2 Confidentiality

Suppose now that the policy maker's information is confidential. The lobbyist is aware of the investigation commissioned by the policy maker (p), but does not know the conclusions of this investigation (r).

[17] The policy maker does not gain from the lobbyist's information because it never makes her change her policy choice. She either continues to prefer policy $x = 0$ or is indifferent between the two policies.

Lobbyist's persuasion strategy

Under confidentiality, the lobbyist's persuasion strategy is based on his beliefs about the realisation of the policy maker's investigation. Each of these realisations defines a *type* of the policy maker which I denote by the realisation of the signal r.

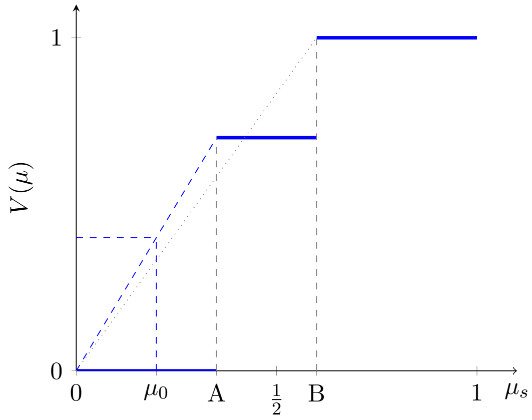
When the policy maker's interim beliefs (μ^r) are always below $\frac{1}{2}$, one realisation of the lobbyist's strategy (s_0) will not persuade any type. The lobbyist needs to choose between generating favourable evidence (s_1) that persuades both the *sceptical* type r_0 and the *sympathetic* type r_1 and favourable evidence that only persuades the sympathetic type r_1.

When the policy maker's own preliminary information is sometimes persuasive ($\mu^{r_1} \geq \frac{1}{2}$), the favourable realisation (s_1) should always persuade both types but the lobbyist chooses whether the unfavourable realisation (s_0) should fully reveal the state to be $\omega = 0$ as before, or if it should be sufficiently imprecise that the sympathetic type (r_1) still prefers policy $x = 1$, i.e. $\mu^{r_1}_{s_0} \geq \frac{1}{2}$.

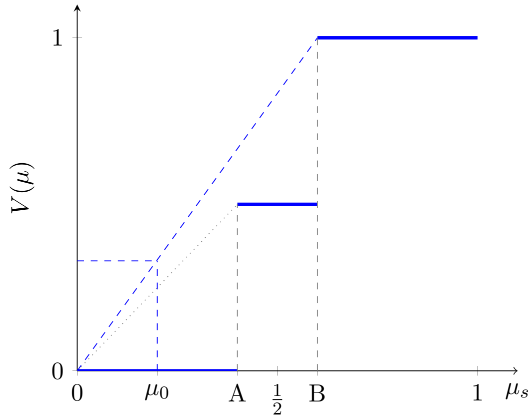
In equilibrium, the lobbyist will choose one of two persuasion strategies. In particular, restricting the lobbyist's persuasion strategy to binary signals is without loss of generality. This is illustrated in figure 1.3 for the case where the policy maker is never persuaded by her own information ($\bar{\mu} < \frac{1}{2}$). If the lobbyist's favourable evidence is not very informative, and only induces a belief at point A, it only persuades the sympathetic type (μ^{r_1}), and generates a payoff less than 1. If it is more informative, and induces a belief at B, it can persuade both types and generates a payoff of 1. The optimal persuasion strategy can be determined by the concavification of this step function.[18]

I refer to the optimal persuasion strategy which targets the sympathetic type r_1 as the *targeted* persuasion strategy and denote it π_T. If the sympathetic type is not already persuaded ($\mu^{r_1} < \frac{1}{2}$), this persuasion strategy should be designed as if the lobbyist knew that the policy maker had observed r_1, that is $\pi_T = \pi_{r_1}$, as defined in Lemma 1. When the sympathetic type is already persuaded ($\mu^{r_1} > \frac{1}{2}$), favourable evidence (s_1) should persuade the sceptical type ($\mu^{r_0}_{s_1} = \frac{1}{2}$) and unfavourable evidence (s_0) should leave the sympathetic type just persuaded to choose policy $x = 1$ ($\mu^{r_1}_{s_0} = \frac{1}{2}$).

[18] Kolotilin (2018) shows that the problem faced by the lobbyist allocating probabilities across different possible realisations is a linear programming problem. The relative marginal gains and marginal costs associated with each realisation can be ranked and one of of the two possible persuasive realisation will dominate the other. The lobbyist therefore always prefers either the persuasion strategy that exactly persuades the favourable type r_1 or the one that persuades both, but never a combination of these strategies.



(a) Persuade only sympathetic type



(b) Persuade both types

Figure 1.3: In equilibrium, the lobbyist only chooses one of two strategies with binary signals

Definition 1. ***Targeted persuasion strategy.*** *π_T is the persuasion strategy defined by:*

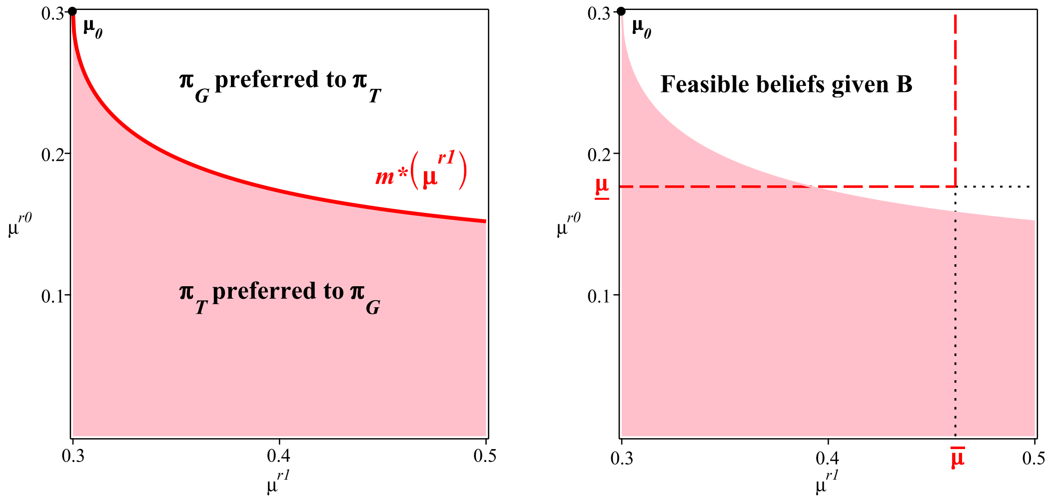
$$\pi_T \ s.t. \ \begin{cases} \pi_T(s_1|1) = 1 \ \ and \ \pi_T(s_1|0) = \frac{\mu^{r_1}}{(1-\mu^{r_1})} & if \ \mu^{r_1} < \frac{1}{2} \\ \frac{\pi_T(s_1|0)}{\pi_T(s_1|1)} = \frac{\mu^{r_0}}{(1-\mu^{r_0})} \ \ and \ \frac{\pi_T(s_0|0)}{\pi_T(s_0|1)} = \frac{\mu^{r_1}}{(1-\mu^{r_1})} & if \ \mu^{r_1} \geq \frac{1}{2} \end{cases}$$

I call the optimal persuasion strategy which persuades both types a *general* persuasion strategy and denote it π_G. This strategy should target the sceptical type r_0 and be designed as if the lobbyist knew that the policy maker had observed r_0: $\pi_G = \pi_{r_0}$, as defined in Lemma 1.

Definition 2. ***General persuasion strategy.*** *π_G is the persuasion strategy defined by:*

$$\pi_G(s_1|1) = 1 \ \ and \ \pi_G(s_1|0) = \frac{\mu^{r_0}}{(1-\mu^{r_0})}$$

Which of these two strategies is optimal depends on the relative likelihood of the two types (captured by the height of the first step in figure 1.3) and the relative distance between the beliefs (captured by the distance between points A and B). Specifically, a general strategy π_G requires more favourable evidence, and makes the favourable results (s_1) less likely to arise. On the other hand, persuading only the more sympathetic type with a targeted strategy π_T is easier (requires less evidence and makes s_1 more likely), but the lobbyist no longer guarantees that favourable evidence (s_1) always persuades the policy maker. A similar intuition applies when the policy maker's own preliminary information



(a) Incentive constraint

(b) Expertise constraint

Figure 1.4: Set of interim beliefs and incentive and expertise constraints

is sometimes persuasive ($\mu^{r_1} \geq \frac{1}{2}$). Formally, the lobbyist chooses π_G over π_T if:

$$\mathbb{P}_{\pi_G}(s_1) > \mathbb{P}_p(r_1)\,\mathbb{P}_{\pi_T}(s_1|r_1) \text{ if } \mu^{r_1} < \frac{1}{2}$$
$$\text{and} \quad \mathbb{P}_{\pi_G}(s_1) > \mathbb{P}_{\pi_T}(s_1) + \mathbb{P}_p(r_1)\,\mathbb{P}_{\pi_T}(s_0|r_1) \text{ if } \mu^{r_1} \geq \frac{1}{2}$$

These conditions can be expressed entirely in terms of the interim beliefs of the policy maker, μ^{r_0} and μ^{r_1}. As the sceptical type's belief becomes more sceptical (μ^{r_0} decreases), the policy maker becomes more likely to be sympathetic ($r = r_1$), and a targeted strategy π_T becomes more attractive. As a result, the lobbyist chooses the targeted strategy π_T if the sceptical type is sufficiently sceptical and the general strategy otherwise. This leads to the following characterisation of the lobbyist's equilibrium persuasion strategy.

Lemma 2. *Under confidentiality, there is a threshold $m^*(\mu^{r_1}) \in (0, \mu_0)$ such that, for any $\mu^{r_1} \in (\mu_0, 1)$, the lobbyist chooses a general persuasion strategy π_G if the belief μ^{r_0} is above $m^*(\mu^{r_1})$, and chooses a targeted strategy π_T otherwise.*

I refer to the condition $\mu^{r_0} \geq m^*(\mu^{r_1})$ as the *incentive constraint.* By contrast, I call *expertise constraint* the bounds imposed on μ^{r_0} and μ^{r_1} by the policy maker's expertise: $\mu^{r_0} \geq \underline{\mu}$ and $\mu^{r_1} \leq \bar{\mu}$, and say that beliefs are *feasible* if they satisfy the expertise constraint. These constraints are illustrated in figure 1.4.

Intuitively, there are two reasons why the lobbyist prefers a targeted persuasion strategy π_T if the interim belief of the sceptical type r_0 is too low. First, the more sceptical the policy maker is, the more evidence the lobbyist has to provide with a general persuasion strategy. Second, the more sceptical the policy maker is, conditional on observing r_0, the more likely it is ex-ante that the policy maker is sympathetic ($r = r_1$). Conversely, as the belief of the sympathetic policy maker μ^{r_1} increases, the probability of r_1 decreases, which leads the lobbyist to prefer the general persuasion strategy π_G. The threshold $m^*(\mu^{r_1})$ therefore decreases as the sympathetic type becomes more sympathetic (μ^{r_1} increases).[19]

Policy maker's preliminary investigation

I now turn to the policy maker's choice of preliminary investigation. I first show under what conditions the policy maker strictly benefits from confidential information. I then show that these benefits are limited by the lobbyist's incentive constraint and characterise the policy maker's optimal choice of preliminary investigation.

Gains from confidentiality When information is confidential, the policy maker's expected utility as a function of her interim beliefs is

$$U^C(\mu^{r_0}, \mu^{r_1}) = \begin{cases} \sum_{r\in\{r_0,r_1\}} \mathbb{P}(r) \sum_{s\in\{s_0,s_1\}} \mathbb{P}_{\pi_G}(s|r) U\left(\mu_s^r\right) & \text{if } \mu^{r_0} \geq m^*(\mu^{r_1}) \\ \sum_{r\in\{r_0,r_1\}} \mathbb{P}(r) \sum_{s\in\{s_0,s_1\}} \mathbb{P}_{\pi_T}(s|r) U\left(\mu_s^r\right) & \text{if } \mu^{r_0} < m^*(\mu^{r_1}) \end{cases} \tag{1.2}$$

Where, as before, $\mathbb{P}(r)$, is determined by the pair of interim beliefs (μ^{r_0}, μ^{r_1}) and the Bayes plausibility constraint: $\mathbb{P}(r_0)\mu^{r_0} + \mathbb{P}(r_1)\mu^{r_1} = \mu_0$.

The policy maker's gain from confidential information arises when the lobbyist chooses a general persuasion strategy: π_G is designed to persuade the sceptical type r_0 who requires more evidence to be persuaded. This additional evidence is valuable to the sympathetic type r_1 who gains some informational rent. In other words, had the lobbyist known that the policy maker was sympathetic, he would have provided less evidence.

The rent obtained from confidential information is captured by the distance between the belief that the lobbyist would like to induce following favourable evidence ($\frac{1}{2}$), and the belief the sympathetic type actually has ($\mu_{s_1}^{r_1}$). When the latter is strictly above $\frac{1}{2}$, the policy maker is more confident that choosing policy $x = 1$ is the correct thing to do: she is better off because her uncertainty is reduced. This intuition is illustrated in figure 1.5.

When the lobbyist chooses a targeted strategy π_T, the policy maker gains no informational rent since the beliefs induced by the lobbyist never make her strictly prefer policy

[19] In fact, an opposite effect also takes place: as μ^{r_1} increases, the type r_1 becomes easier to persuade which makes the lobbyist less likely to choose π_G. The first effect dominates since the decrease in the probability of the type being r_1 is convex, while the gain in the probability of realisation s_1 given r_1 is linear.

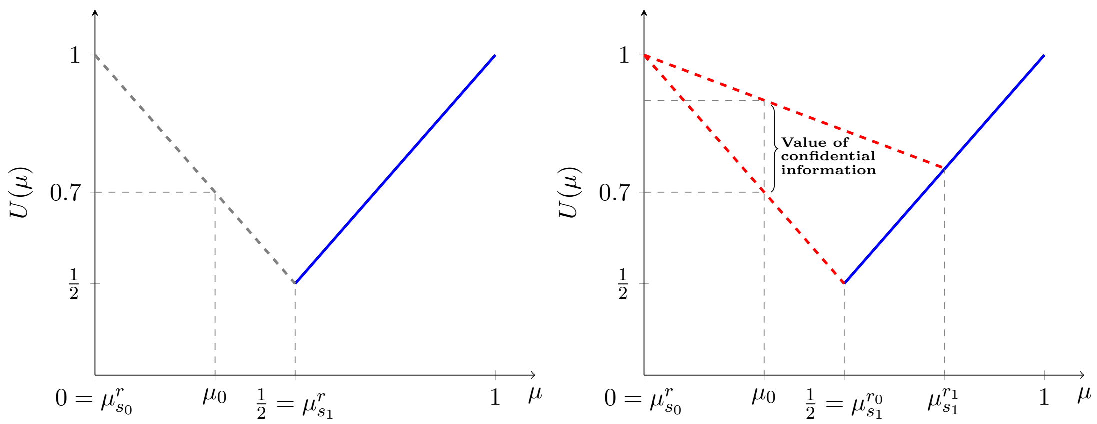


Figure 1.5: Distribution of posterior beliefs when the policy maker's information is public (left) and when it is confidential and the lobbyist chooses π_G (right).

$x = 1$. A targeted persuasion strategy π_T therefore yields the same payoff as public information.[20]

These results are formalised in Proposition 2.

Proposition 2. *For any pair of interim beliefs (μ^{r_0}, μ^{r_1}) the policy maker strictly gains from confidentiality, $U^C(\mu^{r_0}, \mu^{r_1}) > U^P(\mu^{r_0}, \mu^{r_1})$, if and only if (μ^{r_0}, μ^{r_1}) satisfies the incentive constraint: $\mu^{r_0} \geq m^*(\mu^{r_1})$.*

For low levels of expertise, the incentive constraint is satisfied even when the policy maker uses the most informative investigation $\bar{p}$, and the policy maker strictly gains from confidentiality relative to the optimal investigation under transparency. Interestingly, if expertise is so low that $\bar{\mu} < \frac{1}{2}$, the policy maker would not change her decision based on her own information and her information would have no value in the absence of lobbying. The policy maker therefore does not gain from her own information because that information helps her decision *directly*, or because she uses this information to *audit* the information provided by the lobbyist.[21] Instead, the policy maker can gain from her information *indirectly*, by inducing the lobbyist to provide additional evidence. While it is natural to expect the policy maker to gain from having a second source of information, this mechanism allows her to gain even if that second source of information is redundant from a policy choice perspective.[22]

[20] When information is public, the policy maker allows the lobbyist to perfectly target each of her types. Since the targeted strategy is designed to just persuade the sympathetic policy maker, a sceptical policy maker obtains little information. By revealing herself to be sceptical, the policy maker would force the lobbyist to provide additional evidence to persuade her. However, that additional information would never be sufficient to make her change her policy choice, so she does not strictly gain from it.

[21] As for instance in Dellis & Oak (2018).

[22] The information is redundant in the sense that the lobbyist's strategy will reveal at least as much information as the policy maker's investigation. Alonso & Câmara (2018) also show that strategic uses of redundant information can arise when the sender (rather than the receiver) is privately informed.

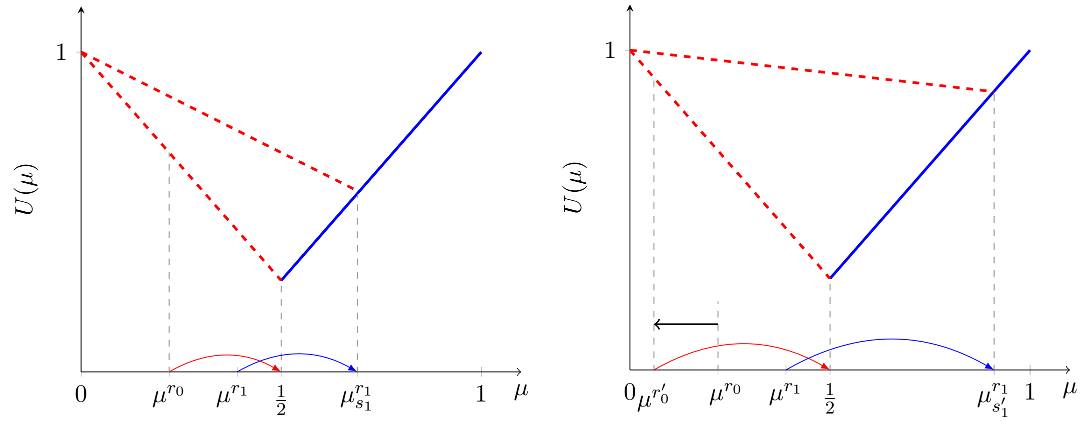


Figure 1.6: Reducing μ^{r_0} to $\mu^{r_0'}$ forces the lobbyist to provide more evidence to induce $\mu^{r_0}_{s_1} = \frac{1}{2}$. This increases $\mu^{r_1}_{s_1}$ and the policy maker's expected utility.

Limits of confidentiality and optimal investigation Given a general persuasion strategy, the policy maker prefers to be as sceptical as possible as this forces the lobbyist to provide more evidence. This is illustrated in figure 1.6: the lower the sceptical belief μ^{r_0}, the more evidence the lobbyist needs to produce under a general strategy π_G to ensure that the sceptical policy maker chooses policy $x = 1$ following a good signal ($\mu^{r_0}_{s_1} = \frac{1}{2}$), and the higher the posterior belief of the sympathetic policy maker following a good signal s_1. The policy maker's expected utility given π_G is therefore decreasing in μ^{r_0}.[23]

However, to induce the lobbyist to choose that strategy, the policy maker may need to distort her information. When expertise B is high, making full use of expertise (choosing $p = \bar{p}$) induces beliefs $(\underline{\mu}, \bar{\mu})$ that violate the incentive constraint $(\underline{\mu} < m^*(\bar{\mu}))$. The lobbyist therefore chooses a targeted strategy π_T (Lemma 2) which provides less evidence than a general strategy π_G.

The policy maker therefore faces a trade-off: to *extract* more information from the lobbyist, she needs to distort the preliminary information she *obtains*. The distortion needs to ensure that the sceptical type is sufficiently likely and not too hard to persuade, so the policy maker chooses a preliminary investigation that makes her sceptical type not too sceptical ($\mu^{r_0} \geq m^*(\mu^{r_1})$). This limits the value of confidential information.

These observations lead to the following characterisation of the policy maker's optimal preliminary investigation.

Proposition 3. *Under confidentiality, there exist thresholds $\underline{B}$ and $\overline{B}$ on the policy maker's expertise such that:*

[23]Note that the policy maker's expected utility is independent of μ^{r_1} because the increase in utility due to a higher posterior belief is exactly offset by the lower probability of that belief occurring. However, when the incentive constraint binds $\mu^{r_0} = m^*(\mu^{r_1})$, it is still preferable to choose the highest sympathetic belief $\mu^{r_1} = \bar{\mu}$ since this loosens the constraint: $m^*(\bar{\mu}) < m^*(\mu^{r_1})$, for any $\mu^{r_1} < \bar{\mu}$.

- *The policy maker chooses the most informative preliminary investigation $\bar{p}$ if either $B < \underline{B}$ or $B > \overline{B}$*

- *She imposes distortions on her preliminary investigation if $\underline{B} < B < \overline{B}$ and sets*

$$\frac{p(r_0|0)}{p(r_0|1)} = \frac{1 - m^*(\bar{\mu})}{m^*(\bar{\mu})} \cdot \frac{\mu_0}{1-\mu_0} \quad \textit{and} \quad \frac{p(r_1|1)}{p(r_1|0)} = B$$

The precision of the policy maker's preliminary investigation relative to her expertise B is therefore non-linear in expertise: for low levels of expertise, the policy maker chooses the most informative preliminary investigation, for intermediate levels of expertise she chooses to restrict her information in order to induce the lobbyist to choose a general persuasion strategy, and for higher levels of expertise, she chooses the most informative investigation available again. This is illustrated in figure 1.7. The solid line represents the sceptical type's belief μ^{r_0} induced by the policy maker's equilibrium investigation. The dashed line represents the lowest possible belief $\underline{\mu}$ and the dash–dot line the incentive constraint $m^*(\bar{\mu})$, both as functions of B.

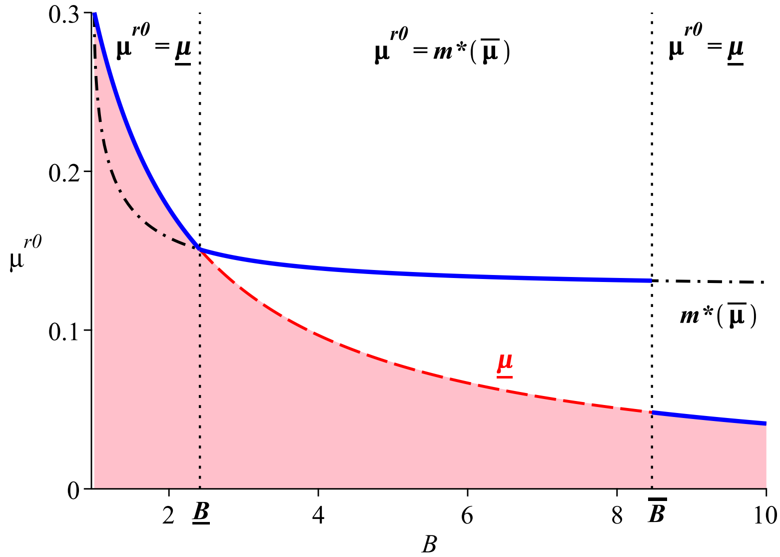


Figure 1.7: Sceptical belief μ^{r_0} induced in equilibrium as a function of expertise B

When expertise (B) is low, the expertise constraint ($\mu^{r_0} \geq \underline{\mu}$) binds before the incentive constraint ($\mu^{r_0} \geq m^*(\bar{\mu})$). The policy maker does not need to distort her information to induce the lobbyist to choose π_G and therefore chooses the lowest sceptical belief possible: $\mu^{r_0} = \underline{\mu}$.[24] The solid line (equilibrium μ^{r_0}) therefore coincides with the dashed line (the

[24]Since her payoff is independent of μ^{r_1}, she can choose any μ^{r_1} such that (μ^{r_0}, μ^{r_1}) is in the set of beliefs that induces π_G. In particular, choosing the most informative investigation is an equilibrium.

lowest possible belief $\underline{\mu}$).

When expertise (B) is intermediate, the incentive constraint binds and the policy maker faces a trade-off between obtaining more precise preliminary information (setting $\mu^{r_0} = \underline{\mu}$) and inducing the lobbyist to provide more evidence (choosing π_G). The loss in expected utility from distorting her information (setting $\mu^{r_0} = m^*(\bar{\mu}) > \underline{\mu}$) is relatively small compared to the gain from extracting information from the lobbyist (inducing π_G instead of π_T). Conditional on inducing the lobbyist to play a general persuasion strategy π_G, she chooses her investigation to induce the lowest possible sceptical belief: $\mu^{r_0} = m^*(\mu^{r_1})$. Since the constraint $m^*(\mu^{r_1})$ is decreasing in the sympathetic type's belief μ^{r_1}, it is optimal to induce $\mu^{r_1} = \bar{\mu}$. The solid line therefore coincides with the dash–dot line (the incentive constraint $m^*(\bar{\mu})$).

Eventually, as expertise (B) becomes sufficiently large, the policy maker may be willing to give up the gains from a general persuasion strategy π_G if the gains from making full use of her expertise ($\mu^{r_0} = \underline{\mu}$ and $\mu^{r_1} = \bar{\mu}$) instead of distorting her investigation ($\mu^{r_0} = m^*(\bar{\mu})$ and $\mu^{r_1} = \bar{\mu}$) are sufficiently large. The solid line (equilibrium μ^{r_0}) thus coincides again with the dashed line (lowest possible belief $\underline{\mu}$). Intuitively, in the limit (as B becomes very large), the policy maker can learn the state almost perfectly and the role of the lobbyist's signal becomes negligible. The policy maker is then willing to give up the gains from inducing the general persuasion strategy π_G in order to use a more precise preliminary investigation.

1.4 The value of confidentiality and its effect on influence

Confidentiality is valuable to the policy maker, but using it to extract information from the lobbyist may require distorting her preliminary investigation. In this section, I show that these distortions can sometimes make confidentiality relatively less attractive. As these distortions depend on government expertise and ideological alignment, I show that the equilibrium value of confidentiality is highest when expertise or alignment is intermediate. Confidentiality also affects how often the lobbyist's preferred policy is passed. I show that the lobbyist's influence also depends non-monotonically on expertise and alignment.

1.4.1 Value of confidentiality to policy maker

The equilibrium value of confidentiality, $W(B, \mu_0)$, is the difference in the policy maker's equilibrium expected utility when information is confidential (U^C), given by expression

(1.2) and when information is public (U^P), given by expression (1.1)

$$W(B, \mu_0) = U^C(m_0, m_1) - U^P(n_0, n_1) \tag{1.3}$$

where m_0 and n_0 are the beliefs of the sceptical type in the confidentiality and transparency equilibria, respectively, and m_1 and n_1 are the beliefs of the sympathetic type in these equilibria.

Value of confidentiality, government expertise, and ideological alignment

The next result reveals that the value of confidentiality varies non-monotonically in both expertise (B) and ideological alignment (μ_0). It is highest when both of these parameters take intermediate values. This is illustrated in figure 1.8 and formalised in Proposition 4 below.

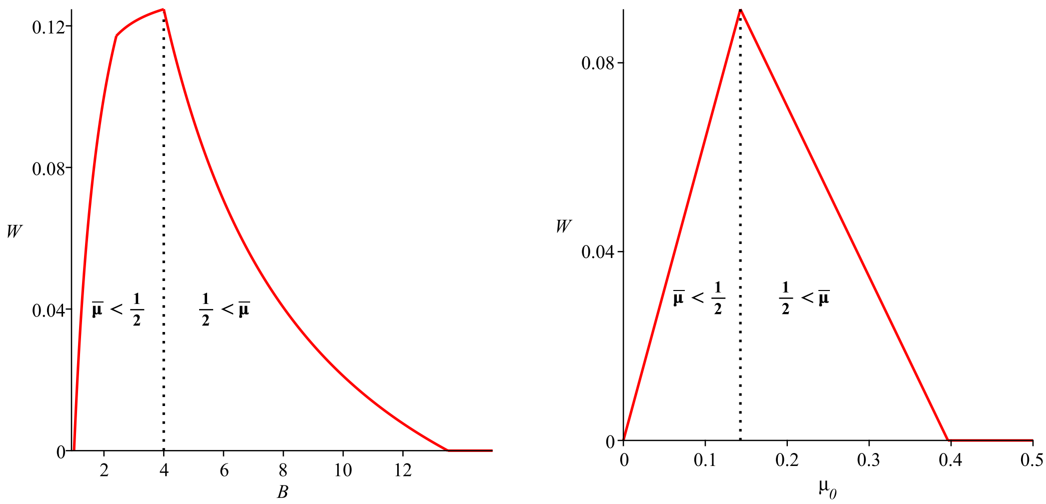
Proposition 4. *The value of confidentiality is*

- *increasing in expertise (B) and in ideological alignment (μ_0) at low levels,*
- *decreasing in both variables at higher levels.*

For sufficiently high expertise or alignment, the policy maker is indifferent between transparency and confidentiality.

As the expertise of the policy maker changes, two opposite effects arise. On the one hand, the value of confidentiality increases because the policy maker can extract more information from the lobbyist when keeping her information confidential: she can make her sceptical belief (μ^{r_0}) more sceptical and force the lobbyist to produce more evidence when choosing a general persuasion strategy π_G. On the other hand, expertise can also increase the value of information when it is public, which decreases the value of confidentiality. Indeed, when expertise is sufficiently high (so that $\bar{\mu} > \frac{1}{2}$), preliminary information is valuable even when public, as the policy maker can base her policy decision on information that the lobbyist would not have provided. That value increases in expertise.

Since the second effect does not arise when expertise is low, the first effect initially dominates and the value of confidentiality increases in expertise. When expertise is high, expertise starts to increase utility under transparency. In addition, the first effect (expertise increasing utility under confidentiality) dampens as the policy maker needs to distort her information (set $\mu^{r_0} = m^*(\bar{\mu})$) to induce the lobbyist to choose a general persuasion



(a) Value of confidentiality as a function of expertise

(b) Value of confidentiality as a function of alignment

Figure 1.8: Value of confidentiality with respect to expertise B and ideological alignment μ_0

strategy π_G. As a result, the effect of expertise on utility is higher under transparency than confidentiality and the value of confidentiality decreases in expertise.

A similar intuition explains the non-monotonicity in ideological alignment. As ideological alignment increases, the policy maker's expected payoff decreases under both transparency and confidentiality. The closer the policy maker's belief is to $\frac{1}{2}$, the easier it is for the lobbyist to persuade her and the less evidence needs to be produced. However, under confidentiality, this effect is mitigated by the policy maker's ability to extract more information from the lobbyist. The policy maker's expected utility therefore decreases with alignment at a slower rate under confidentiality than under transparency, and the value of confidentiality increases in alignment (μ_0) when alignment is sufficiently low ($\bar{\mu} < \frac{1}{2}$).

When alignment is higher, it is possible for the policy maker to benefit from her preliminary investigation when information is public (as $\bar{\mu} > \frac{1}{2}$ is now possible for a given B). The more closely aligned she is to the lobbyist, the more valuable that preliminary information becomes under transparency.[25] Under confidentiality, the policy maker's expected utility continues to decrease in alignment. The value of confidentiality therefore unambiguously decreases when alignment is sufficiently high.

Finally, when expertise or alignment are very large, the policy maker becomes indifferent between confidentiality and transparency. Recall that the policy maker may find distorting her information so demanding that she prefers to let the lobbyist target her

[25]The higher the alignment, the more likely the policy maker is to observe some information (r_1) which makes her more confident about choosing $x = 1$ than she would be with the lobbyist's information (as $\mu^{r_1} > \frac{1}{2}$).

sympathetic type (choose π_T) to be able to make full use of her expertise (Proposition 3). In addition, recall that a targeted strategy yields the same utility as transparency (Proposition 2). The policy maker is therefore happy to make her information public whenever her expertise or alignment is sufficiently high that she would prefer to let the lobbyist choose a targeted persuasion strategy.[26] In fact, in this case, the policy maker would receive strictly less information under confidentiality than under transparency.[27]

Proposition 4 reveals that the choice to make internal information transparent depends on both the policy environment and the political environment.

Given some exogenous value of transparency, transparent institutions should be more prevalent in areas where the government is either not very competent, or on the contrary, very good at obtaining precise policy-relevant information. This depends for instance on the policy's complexity, on whether the government is composed of technocrats, or on whether the civil service is relatively less attractive than the private sector to competent researchers.

Similarly, we should expect more transparency when the policy maker is so opposed to the lobbyist that she does not expect to gain much from the lobbyist's information, or on the contrary, when the policy maker and the lobbyist are so aligned that the policy maker cannot extract much information from the lobbyist.

Value of transparency with weak institutions

How valuable is confidentiality when the policy maker cannot credibly control internal investigations? It may not always be possible for the policy maker to keep the realisations r from her preliminary investigation confidential, yet make the choice of preliminary investigation p public. This possibility affects whether the policy maker can induce the lobbyist to choose a general persuasion strategy π_G and thus benefit from confidential information. If the policy maker's preliminary investigation p is not observable, the lobbyist best responds to the preliminary investigation that she expects the policy maker to choose in equilibrium.

The policy maker would like to commit to choose a preliminary investigation p that induces π_G. However, given that the lobbyist chooses a persuasion strategy π_G, the policy maker would want to deviate to the most informative preliminary investigation, as it would provide her additional information.

[26] This can only occur when the incentive constraint is binding, so high alignment alone is not sufficient.

[27] The policy maker is therefore only indifferent between confidentiality and transparency because the additional information available under transparency has no effect on her policy choice. If she also cared about the total amount of information received, she would strictly prefer transparency over confidentiality.

If expertise is sufficiently low that the lobbyist would still choose a general persuasion strategy (π_G), then the policy maker faces no commitment problem. Otherwise, there can be no equilibrium in which the lobbyist plays a general persuasion strategy π_G. As a result, any equilibrium must involve a targeted persuasion strategy π_T when expertise is large. As shown in Proposition 4, the policy maker is indifferent between transparency and confidentiality when the lobbyist chooses a targeted persuasion strategy π_T. We therefore get the following result.

Proposition 5. *If expertise is sufficiently large, $B > \underline{B}$, and the investigation p is not observed by the lobbyist, the policy maker is indifferent between transparency and confidentiality.*

This result qualifies the finding of the previous section: confidentiality only increases total information available when expertise B is not too large or when the policy maker can credibly commit to distorting her information. This commitment is more credible when institutions are strong: policy makers interact repeatedly with special interest groups, have strong incentives to choose the right policy or can delegate information gathering to independent agencies in the civil service. As a corollary, transparency is beneficial when institutions are weak, and the policy maker cannot commit to distorting preliminary investigations, but only if expertise is large.

1.4.2 Effect of confidential information on influence

In this section, I analyse how the policy maker's control over her investigation affects the lobbyist's influence. I show that an increase in government expertise can sometimes result in both higher welfare for the policy maker and higher influence by the lobbyist, while a decrease in ideological alignment can increase both the policy maker's expected payoff and the lobbyist's influence. This result cautions against the popular view that external influence always has a negative impact on policy making.

Evaluating influence

I define influence as the effect that the presence of a lobbyist has on policy choice. In particular, since the lobbyist's objective is to persuade the policy maker to enact policy $x = 1$, influence is measured as the difference in the ex-ante probability that the policy will be $x = 1$ with and without the lobbyist. This measure is therefore related to the lobbyist's expected utility, which is equal to the probability that the policy chosen is $x = 1$. It differs because it accounts for the policy that the policy maker would choose in the absence of lobbying. Explicitly modeling the policy maker's choice of information therefore highlights that her equilibrium strategy in the counterfactual (the policy choice in the absence of lobbyist) may be different.[28]

In the absence of lobbying, the policy maker weakly prefers the most informative preliminary investigation. The probability of choosing policy $x = 1$ is therefore 0 if the policy maker's expertise is too low to ever change her choice ($\bar{\mu} < \frac{1}{2}$) and it is equal to the probability of observing signal r_1 otherwise.

In the presence of lobbying, the lobbyist chooses a general persuasion strategy π_G in equilibrium, when the policy maker prefers information to be confidential ($B < \overline{B}$). The probability of inducing policy $x = 1$ is therefore the probability of producing a signal $s = s_1$. When the policy maker prefers information to be public ($B > \overline{B}$), the probability that the policy is $x = 1$ is also the probability that $s = s_1$, unless the policy maker is already persuaded ($r = r_1$ and $\bar{\mu} > \frac{1}{2}$).

Thus, influence is measured as

$$F(B, \mu_0) = \begin{cases} \mathbb{P}_{\pi_G}(s_1) & \text{if } \bar{\mu} < \frac{1}{2} \\ \mathbb{P}_{\pi_G}(s_1) - \mathbb{P}_p(r_1) & \text{if } \bar{\mu} \geq \frac{1}{2} \text{ and info confidential} \\ \left[\mathbb{P}_p(r_0)\,\mathbb{P}_{\pi_{r_0}}(s_1|r_0) + \mathbb{P}_p(r_1)\right] - \mathbb{P}_p(r_1) & \text{if } \bar{\mu} \geq \frac{1}{2} \text{ and info public.} \end{cases} \tag{1.4}$$

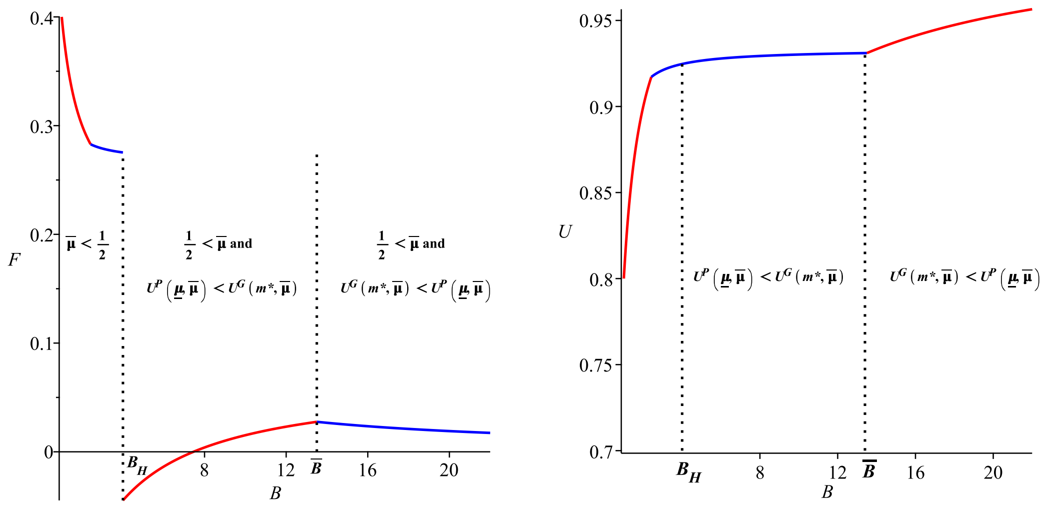
Influence, government expertise, and ideological alignment

I now show how this measure of influence varies with expertise and alignment. Different values of expertise and alignment lead to different combinations of equilibrium strategies with and without the lobbyist. Under some combinations, influence can increase in expertise and alignment, while under others it can decrease in both parameters.

Proposition 6. *The lobbyist's influence on policy making is non-monotonic in the policy maker's expertise (B) and in ideological alignment (μ_0).*

As expertise increases, the policy maker is better equipped to defend herself. She can make the lobbyist believe that she is very sceptical and force him to produce a large amount of evidence. Therefore, as expertise B increases, influence initially decreases. When expertise becomes too large and the policy maker needs to distort her preliminary investigation, this effect is dampened and influence decreases slower.

[28] For a similar argument based on the role of outside lobbying, see Wolton (2018).



(a) Influence as a function of expertise B (b) Welfare as a function of expertise B

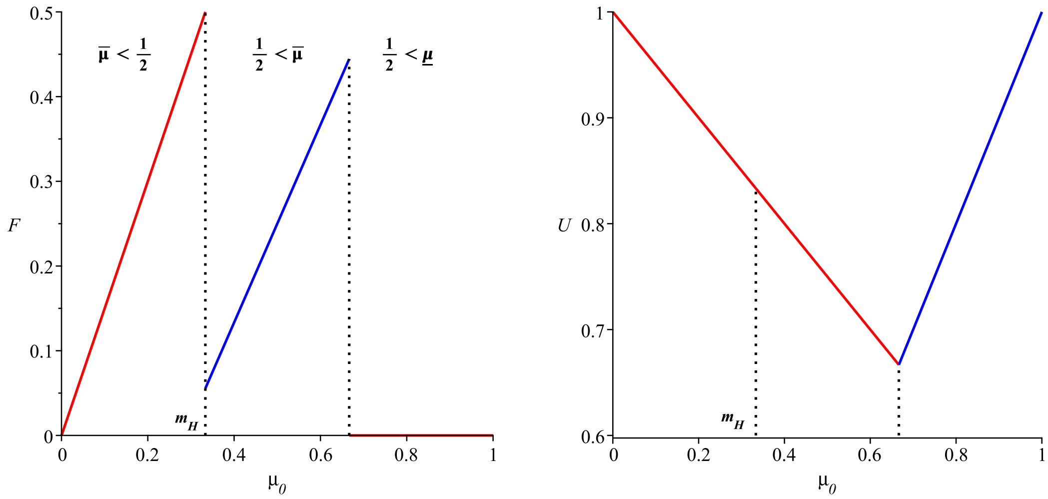
Figure 1.9: Influence, welfare and expertise B

However, influence can also increase in expertise. Since $\omega = 0$ is more likely than $\omega = 1$ ($\mu_0 < \frac{1}{2}$), an increase in expertise makes the policy maker's signal more precise, and her preliminary investigation is more likely to indicate that the state is $\omega = 0$. In the absence of lobbying, the policy maker therefore becomes more likely to choose policy $x = 0$. By contrast, the presence of a lobbyist leads to a relatively high probability that the policy chosen is $x = 1$, especially as the policy maker's ability to extract information becomes more limited. As a result, influence increases in the policy maker's expertise B. Eventually, the policy maker's expertise is so high that she weakly prefers to make her information public. The lobbyist's capacity to influence policy decreases again as the policy maker relies less on the lobbyist's information and more on her own. This non-monotonicity is illustrated in the left-panel of figure 1.9.

Increases in ideological alignment between the policy maker and the lobbyist (μ_0) generally increase influence, but influence drops discontinuously when alignment is large enough that $\bar{\mu} > \frac{1}{2}$ becomes possible.

More ideological alignment leads to more influence for two reasons. First, the most sceptical belief of the policy maker (μ^{r_0}) becomes less sceptical, so it becomes easier for the lobbyist to persuade her to choose policy $x = 1$. Second, it is less likely that the policy maker actually observes a signal from her preliminary investigation that makes her sceptical, so she needs to distort her investigation even more to force the lobbyist to choose a general persuasion strategy π_G.

However, for a given level of expertise, an increase in alignment also means that the



(a) Influence as a function of alignment μ_0 (b) Welfare as a function of alignment μ_0

Figure 1.10: Influence, welfare and alignment μ_0

policy maker's own information becomes relevant: she may decide to enact policy $x = 1$ following a signal r_1 from her own investigation if $\mu^{r_1} > \frac{1}{2}$. At this point, influence drops: the lobbyist still makes the choice of policy $x = 1$ more likely than if there was no lobbying, but the effect of his presence is smaller because the policy maker would have made that choice with some probability based on her own information. As a consequence, it is possible for influence to decrease when alignment increases. These variations are illustrated in the left-panel of figure 1.10.

It is interesting to note that influence can therefore move in the same direction as the policy maker's utility on some ranges. As discussed in section 1.3.2, the policy maker's expected utility is everywhere increasing in her expertise B. If the policy maker is benevolent, i.e. matching the state of the world is the socially beneficial action, then her expected utility identifies social welfare (excluding the lobbyist). Therefore, social welfare increases everywhere in B. This leads to the following result, illustrated in figure 1.9.

Corollary 1. *There exists a range* $\left[B_H, \overline{B}\right]$ *such that both influence and welfare increase in* B *when* $B \in \left[B_H, \overline{B}\right]$.

Since the lobbyist always provides information that would not be otherwise available to the policy maker, lobbying is valuable in this setup. In that sense, it is to be expected that the policy maker's interests and those of the lobbyist may be aligned. However, the presence of lobbying implies that policy $x = 1$ is chosen more often than it would in its absence. Corollary 1 shows that this is not always against the interest of the policy maker, or detrimental to social welfare. In some environments, more influence can be associated with better policy making. As a result, while increasing expertise can also increase influence, this effect should not stop the acquisition of expertise as more expertise always makes the policy maker better-off.

Similarly, influence and welfare can co-move as alignment increases. The policy maker's expected utility is decreasing in alignment, as an increase in alignment corresponds to higher uncertainty and less information provided by the lobbyist. Since influence can also decrease with alignment, we obtain the following result illustrated in figure 1.10.

Corollary 2. *There is a threshold m_H such that influence drops when $\mu_0 = m_H$, and welfare decreases on an interval around m_H.*

The relationship between influence and ideological alignment also highlights a counterintuitive effect of lobbying: while we would expect influence to be highest when the lobbyist and the policy maker are ideologically aligned, Corollary 2 shows that influence might decrease with alignment, because a more aligned policy maker would have obtained favourable information on her own in the absence of lobbying.

Finally, it is interesting to note that influence may be negative as shown in figure 1.9. In other words, the probability of enacting the lobbyist's preferred policy can be higher without the lobbyist than with him. This occurs because without a lobbyist, the policy maker would choose the most informative preliminary investigation whereas when facing a lobbyist, she chooses to distort her information. This distortion increases the probability of the policy maker being sceptical and therefore decreases the probability that she chooses the lobbyist's preferred policy $x = 1$, even following the lobbyist's persuasion attempt. This negative value of influence implies that the lobbyist would like to *commit* not to intervene in the policy process for some levels of expertise and alignment. By distorting her information, the policy maker therefore forces the lobbyist to intervene and provide information.

Corollary 3. *Confidentiality can force the lobbyist to provide information when he would prefer not to intervene ex-ante.*

1.5 Discussion

I first discuss the implications of these results for the measurement of interest group influence, and then for the relationship between transparency of the policy process and the quality of policy making.

Empirical assessment of interest group influence

The results presented in the previous sections have two implications for the interpretation of studies of interest group influence. First, since government expertise affects both the amount of evidence provided by special interest groups and the choice of policy, failing to include expertise can lead to omitted variable bias. Second, interpreting the effect of lobbying on policy change as *influence* relies on the wrong counterfactual. Instead, influence should be assessed relative to the policy that would have been chosen in the absence of lobbying.

Influence is usually measured as the effect of lobbying expenditures on policy changes benefiting special interest groups. For instance, de Figueiredo & Silverman (2006) estimate the return from money spent on lobbying by universities onto the earmarks received by these universities, Richter et al. (2009) analyse the effect of lobbying expenditures by firms on their effective tax rate, Igan & Mishra (2014) look at the return on lobbying expenditures in the financial industry, Kang (2016) focuses on the effect of lobbying resources on the probability of a specific policy being passed within the energy industry, and Payson (2017) looks at the changes in revenues from states to cities induced by these cities' lobbying effort. All these studies find significant effects of lobbying expenditures on policy change, and estimate large returns to these expenditures.

The model presented here abstracted from information costs to emphasise that trade-offs between internal and external information can arise for strategic reasons rather than purely monetary considerations. Assuming instead that costs of producing information are proportional to the precision of information produced (see Gentzkow & Kamenica (2014) for a discussion) would not affect the conclusions of the model, as long as the gains of each player from obtaining their preferred policy are large enough relative to the costs of producing information.

In that case, resources spent on lobbying – which correspond to an increase in the informativeness of the information provided – would be positively related to the policy maker's level of expertise B, since the higher expertise, the more sceptical the policy maker's belief (μ^{r_0}) could be, and the more evidence the lobbyist would have to provide in equilibrium. Similarly, an increase in ideological alignment should be related to a decrease in resources as persuasion becomes easier. Most studies account for ideological alignment, but few control for government expertise.[29] To show how this affects the estimates of

[29] An exception is Igan & Mishra (2014) who interact legislator and Congress fixed effects to account for changes in expertise. This would still fail to capture shorter term changes, however.

influence, consider the following regression equation:

$$x_i = f(\alpha + \beta R_i + \delta M_i) + \varepsilon_i \tag{1.5}$$

Where $x_i \in \{0, 1\}$ is the policy chosen on bill i, R_i are resources spent by lobbyists, and M_i is the alignment between policy makers and lobbyists.

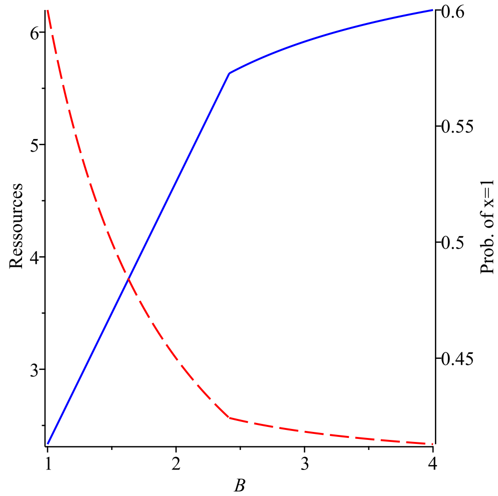


Figure 1.11: Lobbying expenditures (solid line) vs. probability of policy $x = 1$ (dashed line)

If the data includes lobbyists who spend close to no resources and have no influence, we should expect $\beta > 0$.[30] The model shows that a relevant variable is bill-specific government expertise B_i, that this variable is positively correlated with resources, and negatively correlated with the probability of choosing policy $x = 1$ when it is not chosen ex-ante ($\mu_0 < \frac{1}{2}$), as illustrated in figure 1.11.[31] This means that the error term in equation (1.5) is $\varepsilon_i = \gamma B_i + \nu_i$, where $\gamma < 0$ and $\text{cov}(B_i, R_i) > 0$. As a result, estimating equation (1.5) would produce a downwardly biased estimate of β. This can be dealt with by controlling for government expertise, which could be proxied by the budget of agencies dedicated to the relevant policy domain, the number of reports produced internally on a specific policy, or controlled for by policy-level fixed effects if the identification variation allows for it.

The second implication is that measuring influence through x_i, the probability of a

[30] While not explicitly covered by this model, this positive relationship could come from the presence of some lobbies spending resources even when their resources are too constrained to affect the policy choice, i.e. when $\mu_{s_1} < \frac{1}{2}$ for any π. Lobbyists could do this to ensure future access, or to influence other policies as in Ellis & Groll (2017). Lower resources would then be associated to no policy change, while higher resources would correspond to policy change.

[31] For simplicity, resources are assumed to be equal to the precision of signal s_1 chosen in equilibrium by the lobbyist, which is the only precision that varies on this range.

policy change, is based on the wrong counterfactual. As the model shows, the equilibrium investigation of the policy maker is affected by the presence of lobbying. This is important for the interpretation of coefficients. Consider for example the policy implications from a positive estimate of the parameter δ in equation (1.5), the effect of ideological alignment on policy change. If x_i measured influence, this would imply that influence decreases when ideological alignment decreases. However, Proposition 6 shows that when ideological alignment decreases, the policy maker may no longer choose policy $x = 1$ in the absence of lobbying. As a result, the role of the lobbyist becomes more important in changing policy and actual influence may increase.

Evolution of institutions

The results from Section 1.4 provide a rationale for the development of some institutions used for information gathering in legislatures and contributes to a more general literature on the role and development of legislative institutions (see Krehbiel (1991), Bimber (1991), Krehbiel (2004) or Stephenson (2011)).[32] In particular, the observed diversity of systems used to obtain information within governments (agencies, legislative research services, public consultations, legislative hearings, etc.) raises a number of questions: why might legislators obtain redundant information from both internal and external sources? Are internal and external sources substitutes or complements? What explains the transitions from confidentiality to transparency, such as congressional hearings in the U.S. in the 1970s and congressional research memos more recently?[33] The model's predictions can be related to these observations.

First, the model shows that internal information is valuable to the policy maker even when that information would not impact policy in the absence of lobbying (Proposition 2). This can account for the puzzling observation that the government may choose to obtain information even if that information is redundant. For example, one Appropriations committee aide stated that "Congress relies on CRS as an extension of staff, for quick and

[32]Existing studies of information generation in legislatures have focused on the incentives of agents to acquire costly information, and how institutional features, such as voting or agenda setting can affect these incentives. A large literature has developed around the seminal work of Giligan and Krehbiel (Gilligan & Krehbiel (1987), Gilligan & Krehbiel (1989), and Gilligan & Krehbiel (1990)) which shows how decision-making procedures (in particular closed vs. open rule) affect the incentives of legislative committees to acquire and transmit information in the legislature.

[33]In 2018, the US Congress eventually required the congressional research discussed in the introduction to be made publicly available (DeBonis 2017). This change only concerns general reports produced by the CRS and not specific reports requested by legislators with confidentiality requirements. In practice, many of these general reports had become available through connections to individuals with insider access, but the CRS's original publication policy created barriers that limited the ease of access of these reports, which the more recent changes overturned. This policy change echoes the Legislative Reorganization Act 1970, which required open congressional hearings to be televised.

dirty analysis that is sometimes not perfect" (Clark 2016), which suggests that the evidence obtained internally may not be precise enough to determine policy choices. The model suggests that 'quick and dirty analysis' can result in significant policy improvements in the presence of lobbyists. In addition, higher expertise allows policy maker to produce more internal information, and at the same time to induce lobbyists to produce more evidence, so the two sources can be complements.

Second, Proposition 4 shows that for a large enough level of expertise or ideological alignment, confidentiality is no longer valuable, and Proposition 5 shows that this effect is even stronger when the policy maker cannot credibly convey the type of preliminary investigation that she carries out. As a result, increases in expertise within the government or changes in ideological alignment with lobbyists can lead to more transparency of government information. This can arise independently of the role of transparency for accountability (as in Argenziano & Weeds (2017) for example). In addition, since expertise varies across policy areas (Howlett 2015), it is possible for transparent institutions (such as hearings) to be used in certain domains, and confidential ones (such as agency memos) in others even when the ideological alignment between policy makers and special interest groups remains the same.

Finally, since the value of confidentiality decreases with expertise when expertise is high (Proposition 4), we should observe empirically that increases in expertise are associated with increases in transparency. This is consistent with the findings of Islam (2006) that transparency (measured by the timeliness of governments in releasing economic information and the presence of freedom of information laws) is correlated with the quality of governance. However, it would be incorrect to conclude that transparency causes improvements in governance. In particular, Proposition 4 indicates that at low levels of expertise, the value of confidentiality is strictly positive. In other words, transparency would lead to worse policy making in that case.

1.6 Conclusion

This paper examined the effect of a policy maker's internal information on the provision of information by special interest groups. When policy makers can control their preliminary investigations, they can extract additional evidence from special interest groups by distorting these investigations. This gives value to internal information, even when that information is limited. However, this possibility only arises when internal information is kept confidential. This makes confidentiality valuable to policy makers even in the absence of reputational concerns and explains why internal research is kept secret in many governments.

When the government obtains confidential information, special interest groups have to adapt the strategy they use to influence policy. Their ability to affect policy decisions is limited by the uncertainty they face, and they would prefer information to be publicly available. However, the paper also highlights that influence cannot be simply measured based on whether policy changed or not. As the policy maker's own investigation changes in the absence of lobbying, the definition of influence should consider what policy would have been chosen in the absence of lobbying. Given this definition, it is possible that influence and welfare increase at the same time, when government expertise or ideological alignment change. When the Congressional Research Service opens its research to the public, as is currently planned, it will increase the influence that interest groups exert on policy making. But if this move towards transparency is driven by an increase in expertise, then this increase in influence could be accompanied by an increase in welfare.

More generally, the results showed that the value of confidentiality varies with government expertise and with the ideological alignment between policy makers and interest groups. When other factors make transparency more desirable, the model suggests that the tension between transparency and confidentiality will be lowest when government expertise is either high or low. Transparency is therefore more likely to be observed when expertise, and therefore the quality of policy making, is high. Yet, we should not conclude that transparency leads to better policy making. This paper shows that at intermediate levels of expertise and intermediate ideological alignment, imposing transparency would lead to worse policy making.

Understanding the control of governments over the production of internal information is therefore critical not only to the study of special interest groups but also to the relationship between transparency and the quality of government.

Chapter 2

Competition in persuasion between privately informed senders

2.1 Introduction

In many situations, competing agents can influence the behaviour of a decision-maker by providing information. For instance, a pharmaceutical company might carry out tests on one of its drugs to persuade doctors or patients that its product is more effective than those of competitors. Two lobbyists with opposite preferences can produce evidence about the benefits of a policy to influence a policy maker. Two competing media outlets can choose the editorial standards that their investigative journalists should adhere to with the objective to influence the views of their readers.

In all these situations, competing forces can induce these agents to reveal more information. By strategically producing evidence about its products, a firm can indicate to consumers that these products are better than those of its competitor. The competitor cannot take away that information from the consumer, so the only alternative is to generate more evidence about its own product, in the hope that this evidence will persuade the consumer to buy it.

The existing literature (e.g. Gentzkow & Kamenica 2017*b*,*a*) has shown that, as long as senders have access to sufficiently sophisticated technologies to generate information, competition will lead to more information.

This paper shows that this conclusion does not necessarily hold when these senders have private information prior to generating evidence. In that case, more information can be revealed when the senders are merged into one than when they compete.

More specifically, I consider a situation in which two senders compete to persuade a receiver by designing an experiment. The transmission of information operates through Bayesian persuasion: senders choose probability distributions over signal realisations, conditional on an unknown state of the nature, and can commit to revealing these realisations to the receiver. The two senders may also have some private information about the state prior to designing their experiment. Their choice of experiment can therefore signal their private information to the receiver in equilibrium.

When the two senders are merged, two features of their environment change. First, the joint preferences of the senders over the receiver's actions are different than their individual preferences. Second, the senders share their private information. As a result, the merged senders may have different incentives to signal their joint private information than each individual sender. I refer to the situation where the two senders are merged as 'collusion', even though the senders may not be choosing this situation deliberately to increase their payoff.[1]

There are three effects that can result in more information being produced under collusion than competition. The first effect is the signalling role of choosing a given experiment. The same experiment can induce different posterior beliefs, depending on the type of sender who chooses it in equilibrium. When senders are merged, the incentives of different types of the coalition to choose a particular experiment may differ from the incentives of competing individual senders. As a result, the interim beliefs of the receiver given a choice of experiment may be different in collusion than in competition, and the posterior beliefs induced by that experiment will also differ.

The second effect arises through the interaction between the information produced by the experiment chosen by one sender and the private information revealed by the strategy choice of the other sender. Senders cannot 'take away' information that has been revealed by another sender and this restricts the set of distributions over posterior beliefs that senders can induce in competition. In particular, this will occur when a choice of experiment reveals the private information of a sender. Some distributions are therefore impossible to generate in competition but possible in collusion.

Finally, differences between the competition and collusion equilibria can arise because senders are asymmetrically informed. If a sender learns the other sender's private information when the two senders are merged, then the optimal strategy might differ between competition and collusion. If the less informed sender benefits from signalling this private information to the receiver, but the informed sender does not, then the equilibrium in collusion will differ from the equilibrium in competition.

I first identify a set of sufficient conditions on the preferences and information of the senders, under which the least informative collusive equilibrium is more informative than the least informative competitive equilibrium. I focus on the least informative equilibrium since there are typically many equilibria in competition. In particular, full revelation is always an equilibrium: if both senders choose a fully revealing experiment, an individual sender cannot unilaterally undo that revelation.

I then restrict attention to specific utility functions of the senders and show that there exist parameter values such that these sufficient conditions are satisfied. As a result, there exist payoff functions and a private information structure, such that the least informative collusive equilibrium is more informative than the least informative competitive equilibrium. To show this, I focus on a symmetric pooling weak perfect Bayesian equilibria of

[1] I borrow this terminology from Gentzkow & Kamenica (2017*b*).

the two games, and use the intuitive criterion refinement.

In that example, I show that collusion cannot reveal more information than competition if either both senders have perfect knowledge of the state (i.e. observe perfectly revealing signals before designing their experiments), or if they have no information. This indicates that the informativenes ordering of collusion and competition can be non-monotonic in the precision of the senders' private information.

These results have two key implications. First, they reveal that the private information of senders matter to assess whether competition increases the provision of information compared to collusion. While the existing literature has focused on the set of experimenting technologies available to the senders to determine this ordering, I show that another dimension to take into account is the pre-existing information structure.

Second, these results have implications for evaluating the effect of mergers or the design of organisations. The impact of mergers on consumer welfare has traditionally been evaluated by looking at their effect on prices or product quality. However, the effect of mergers on information can also be a relevant concern. For example, the monopoly status of technology conglomerates such as Google or Facebook, whose main service is the provision of information to users, has been criticised.[2] The results in this paper indicate that knowing the information environment of these companies is important to assess the effect of monopoly power on the information they provide to users. In particular, it seems unlikely that these companies can signal any private information about the products they advertise to consumers through their choice of algorithms. As a result, competition would lead to more information generated. By contrast, pharmaceutical companies are more likely to have some information about the effectiveness of their drugs before designing clinical tests. As a result, competition can lead to less information being generated about these products. In addition, the paper also suggests that a manager of a research and development division could decide to force two competing research teams to collaborate, if she is interested in obtaining more evidence about the quality of a project. This would be the case when these teams have done some prior research on the quality of the product. Finally, a legislator could become more informed by commissioning a joint report from opposed interest groups if those have private information regarding the optimality of a policy, but would learn more by hearing their recommendations separately otherwise.

This paper is related to the large literature on information transmission and persuasion,

[2] It has been observed for instance that "Many information monopolies today are more interested in collecting our data than taking our money. The stronger argument is that information monopolies discourage competition, and that ultimately will limit choice and innovation." (see `https://techcrunch.com/2010/11/13/information-monopolies-internet/?guccounter=1`).

and in particular, to models of competition in information transmission. Early works on this question include Milgrom & Roberts (1986) who show that when parties have verifiable information, then competition will induce full revelation of that information as long as one sender prefers the full information outcome. Shin (1998) shows that this result continues to hold even if parties are imperfectly informed, and in particular, even if the receiver is ex-ante as well informed as each sender is. Krishna & Morgan (2001) also show that competition generates more information when informed senders are allowed to publicly send messages sequentially.

The model in this paper uses the Bayesian persuasion approach developed by Kamenica & Gentzkow (2011), and in particular considers the case of Bayesian persuasion by privately informed senders. The case of one individual privately informed sender has been studied by Alonso & Câmara (2018), Hedlund (2017) and Perez-Richet (2014). In all these papers, the choice of an information revelation mechanism can signal the information of the sender. Alonso & Câmara (2018) highlight that a privately informed sender cannot gain and can sometime be made worse-off by their private information. Hedlund (2017) characterises the conditions for pooling or separating equilibria to arise, and Perez-Richet (2014) shows that, in a restricted setting where the sender is perfectly informed, all the sender's types always pool on the same signal. This literature does not consider the case of multiple senders competing against one another.

By contrast, Gentzkow & Kamenica (2017*b*) evaluate the effect of competition on information transmission in the context of Bayesian persuasion but assume that all senders and receivers share a common prior and have no private information. They show that a property of the information environment, that they call Blackwell connectedness, which requires that each sender can unilaterally deviate to a more informative outcome, is necessary and sufficient for competition to be weakly more informative than collusion. They show that when senders compete, they can always deviate to an information revelation strategy that reveals more information, holding the strategy of the other sender constant, but not to one that reveals less information. Gentzkow & Kamenica (2017*a*) focus on the special case where senders have access to any type of signals and show that the result continues to hold. This paper extends these two models by allowing for senders to be privately informed. Boleslavsky & Cotton (2018) also obtain that competition increases information provided when competition is induced by the receiver's limited capacity to implement the recommendation of both senders, so that the receiver can be better off with limited capacity to select projects. In contrast with these papers, Li & Norman (2018) show that additional senders may not increase information when senders cannot arbitrarily correlate their signals to other senders, if senders move sequentially, or if mixed strategies are allowed. In this paper, I show that additional senders may not increase information even when these conditions are not satisfied, provided that senders are privately informed.[3]

The rest of the paper is organised as follows: I present the model and the notation in section 2.2. Section 2.3 provides some generic conditions for the least informative equilibrium in collusion to be more informative than in competition. In section 2.4, I introduce a parameterised problem and derive sufficient conditions in this problem for more information to arise in collusion. Section 2.5 discusses how these results change when the private information of the senders changes and how the welfare of all players is affected. Section 2.6 concludes.

2.2 Model

Environment. There are two senders, I and N, who try to influence a receiver. The receiver's payoff depends on an unknown binary state of the world denoted $\omega \in \Omega = \{L, R\}$. Players have a common prior over the state $\mu_0(\omega) = \mathbb{P}(\omega)$.

Sender I (I for informed) may also receive private signals regarding the state of the world, unobservable by the receiver and the other sender. This private signal is denoted $\theta \in \{\theta_L, \theta_R\}$, and is generated from a distribution $p(.|\omega)$ conditional on the state of the world. The distribution is common knowledge to all players, but not the realisation. I refer to θ as the 'type' of this sender. Following the realisation of this signal, the sender updates his beliefs using Bayes' rule to $\mu_\theta(\omega|\theta) = \frac{p(\theta|\omega)\mu_0(\omega)}{p(\theta|\omega)\mu_0(\omega)+p(\theta|\neg\omega)\mu_0(\neg\omega)}$. I refer to $\mu_\theta := \mu_\theta(R|\theta)$ as the interim belief of a type θ sender.

Preferences. The preferences of the receiver depend on her action a and the state ω and are represented by the utility function $u(a, \omega)$. The preferences of sender $i \in \{I, N\}$ only depend on the receiver's action and are represented by a utility function $v_i(a)$. The

[3]This paper is also related to papers studying competition in information transmission using other approaches than Bayesian persuasion. Martimort & Semenov (2008) use a mechanism design approach to study the incentives of interest groups to collude or to compete and the effect this has on public welfare. Bhattacharya & Mukherjee (2013) study competition between privately-informed senders ('experts') and show that experts with more extreme preferences induce more information to be revealed, but that the effect of private information is ambiguous. In contrast to this paper, their receiver is uncertain about the quality of the sender's information, and the persuasion operates through the disclosure of verifiable information. Finally, Kartik et al. (2017) find that a receiver can be better informed with fewer senders. This arises because the senders' information acquisition choices are substitutes. Their model relies on costly information acquisition rather than the costless information generation of Bayesian persuasion, and their focus is on adding senders rather than comparing collusion and competition.

utility of the senders when they collude is a weighted sum of their individual utilities: $v_m(a) = \alpha v_I(a) + (1-\alpha)v_N(a)$. All payoffs are common knowledge.

Actions. Following Gentzkow & Kamenica (2017*a*), I assume that each sender's strategy, π_i, consists of a finite partition of $\{L,R\}\times[0,1]$, such that $\pi_i \subset S$ where S is the set of non-empty Lebesgue measurable subsets of $\{L,R\}\times[0,1]$. The setup developed by Gentzkow & Kamenica (2017*a*) is then completed by considering a random variable X independent of ω and uniformly distributed on $[0,1]$, and an S-valued random variable equal to s when $(\omega,x)\in s$, for $\omega\in\{L,R\}$ and x, a realisation of X. As a result, given some π_i, the probability of a signal realisation s is: $\mathbb{P}(s|\omega) = \lambda(\{x|(\omega,x)\in s\})$, where $\lambda(.)$ is the Lebesgue measure. Throughout the paper, I refer to these strategies as 'experiments'.

In competition, experiments π_i are chosen simultaneously by both senders and observed by the receiver, and the senders commit to revealing the realisation of this experiment (s) to the receiver. Under collusion, the coalition chooses a unique experiment π, observed by the receiver and also commits to revealing its realisation to the receiver.

Modelling experiments in this way explicitly specifies the joint distribution of the two signal realisations and allow senders to arbitrarily correlate their signal realisations to that of the other senders, thus allowing any Bayes-plausible distribution of posterior beliefs to be induced. Note that, contrary to the case where the two senders generate independent signal realisations, additional signal realisations do not necessarily generate more information in this setup, as one sender's signal realisations can be perfectly correlated with those of the other sender. More details on these rich signal spaces and useful graphical representations of how multiple signals are combined are presented in Gentzkow & Kamenica (2017*a*).

In the competitive game, I denote by $\boldsymbol{\pi}$ the strategy profile of the two senders: $\boldsymbol{\pi} = (\pi_I, \pi_N)$. The information revealed to the receiver is the join of the two experiments: $\pi_I \vee \pi_N$, defined on the lattice induced by the refinement order among experiments (which are partitions of $\{L,R\}\times[0,1]$). In other words, the receiver's information is given by the finer partition of the signal space that arises as each sender's experiment creates additional partition of $\{L,R\}\times[0,1]$. This is illustrated in figure 2.1. In that example, $\pi_I = \{s_L^I, s_R^I\}$ and $\pi_N = \{s_L^N, s_R^N\}$, where

$$s_L^I = (L,[0,0.8]) \cup (R,[0,0.3]) \quad \text{and} \quad s_R^I = (L,[0.8,1]) \cup (R,[0.3,1])$$
$$s_L^N = (L,[0,0.6]) \cup (R,[0,0.4]) \quad \text{and} \quad s_R^N = (L,[0.6,1]) \cup (R,[0.4,1])$$

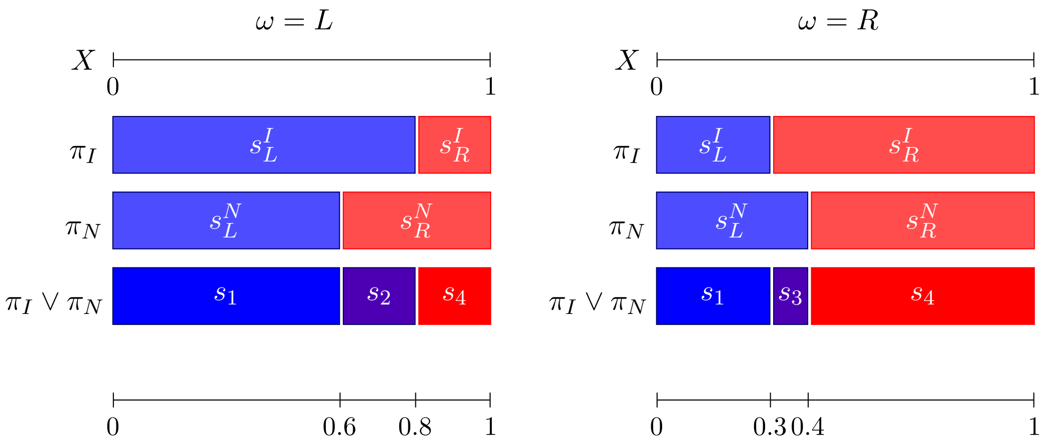


Figure 2.1: Join of two experiments π_I and π_N

As a result, the information gained by the receiver from the two experiments is determined by the realisations: $s_1 = s_L^I \cap s_L^N$, $s_2 = s_L^I \cap s_R^N$, $s_3 = s_R^I \cap s_L^N$, and $s_4 = s_R^I \cap s_R^N$.

The receiver chooses an action $a \in A$, where A is a compact set, after observing the experiments (π_I, π_N) chosen by each sender and a signal realisation s from these experiments. In equilibrium, the receiver updates her beliefs using Bayes' rule. She first forms interim beliefs upon observing the choice of experiments $\boldsymbol{\pi}$, given the equilibrium strategy of a type θ sender $\mathbb{P}(\boldsymbol{\pi}|\theta)$:

$$\mu(\omega|\boldsymbol{\pi}) = \frac{\sum_\theta \mathbb{P}(\boldsymbol{\pi}|\theta)\,\mathbb{P}(\theta|\omega)\mu_0(\omega)}{\sum_{\omega' \in \Omega} \sum_\theta \mathbb{P}(\boldsymbol{\pi}|\theta)\,\mathbb{P}(\theta|\omega')\mu_0(\omega')}$$

Given these interim beliefs, she then updates her posterior beliefs upon observing some signal realisation s:

$$\mu_s(\omega|\boldsymbol{\pi}) = \frac{\mathbb{P}(s|\omega)\mu(\omega|\boldsymbol{\pi})}{\sum_{\omega' \in \Omega} \mathbb{P}(s|\omega')\mu(\omega'|\boldsymbol{\pi})}$$

In the rest of the paper, I use μ to denote the probability of state R ($\mu = \mu(R)$), for prior, interim and posterior beliefs. Since I focus on pure strategies, the possible interim beliefs of the receiver (in equilibrium) are in the set $\{\mu_{\theta_R}, \mu_{\theta_L}, \mu_0\}$.

Timing. The timing is as follows.

1. Nature determines a state of the world, privately from all players

2. Senders receive private signals

   (a) In the competitive game, Sender I receives a private signal $\theta \in \{\theta_L, \theta_R\}$, generated from the distribution $p(.|\omega)$

(b) In the collusive game, the two senders receive that private signal

3. Senders choose their experiments

   (a) In the competitive game, each sender simultaneously chooses an experiment π_i

   (b) In the collusive game, the two senders jointly choose an experiment π

4. The receiver observes the strategies of the senders and the signal realisations, updates her prior and chooses an action

5. Payoffs are realised.

Additional definitions and assumptions

Given a profile of experiments $\boldsymbol{\pi}$, I denote by $\langle \boldsymbol{\pi} \rangle$ the distribution over signal realisations $s \in \boldsymbol{\pi}$ induced by $\boldsymbol{\pi}$, and therefore over posteriors μ_s.

I assume that for a given belief μ of the receiver, there is a unique action that maximises her expected utility. Let $\hat{a}(\mu) = \operatorname{argmax}_{a \in A} \mathbb{E}_{\mu}\left[u(a, \omega)\right]$, that action.

Given that this action is unique, and that the senders' utilities are independent of the state, we can easily map a distribution over posterior beliefs into the expected utility of a sender of type θ. I thus define

$$V_i((\mu_s)_{s \in S}; \mu_r) := \mathbb{E}_{\pi}\left[v_i(\hat{a}(\mu))\right] = \sum_{\omega \in \Omega} \mu_\theta(\omega) \left(\sum_{s \in S} \pi(s|\omega) v_i(\hat{a}(\mu_s)) \right) \tag{2.1}$$

where $(\mu_s)_{s \in S} \in [0,1]^{|S|}$ is the support of posterior beliefs of the receiver induced in equilibrium and μ_r is the receiver's interim belief upon observing the senders' choices of experiments $\boldsymbol{\pi}$. Finally, $\pi(s|\omega)$ is the conditional probability that induces posterior beliefs $(\mu_s)_{s \in S}$ given interim belief μ_r.

Note that if $|S| = 2$, the distribution over posteriors is uniquely pinned down by the Bayes plausibility constraint: $\sum_{s \in S} \mathbb{P}(s) \mu_s = \mu_r$, given some interim belief μ_r of the receiver. Therefore, knowing the support $(\mu_s)_{s \in S}$ and the receiver's interim belief μ_r is enough to compute the sender's expected utility. When $|S| > 2$, there can be multiple distributions satisfying the constraint, but the notation $V_i((\mu_s)_{s \in S}; \mu_r)$ will only be used when it is not ambiguous.

Finally, let $\succsim$ denote the ordering over distributions over posteriors according to Blackwell informativeness.

Equilibrium concept

I characterise weak perfect Bayesian Nash equilibria of the two games. Following an off-equilibrium action π', the receiver forms out-of-equilibrium interim beliefs regarding the type of the deviating sender $\mathbb{P}(\theta|\pi')$. Given these interim beliefs she then updates her beliefs in the same way as she does when facing an on-equilibrium action, based on the signal structure π'.

I abuse notation and refer to these interim beliefs, denoted μ_θ as the 'out-of-equilibrium beliefs' of the receiver. This is used as a shorthand for 'the interim belief that the sender forms about the state of the world following a deviation, given some out-of-equilibrium beliefs about the sender's type following this deviation'.

I use the intuitive criterion (Cho & Kreps 1987) as a refinement on these out-of-equilibrium beliefs with the following modification: if a deviation is either equilibrium dominated for both types, or not equilibrium dominated for either type, then the only possible out-of-equilibrium belief of the receiver is the prior μ_0. In the original version of the intuitive criterion, the receiver is allowed to have *any* out-of-equilibrium beliefs in these situations. This modification rules out equilibria that could be sustained by believing that only one type deviated in these situations, and therefore simplifies the proof. It is also reasonably intuitive insofar as there are no strong argument for assigning weight on one type rather than the other in these cases. The existence result does not rely on this modification.

Preliminary results

Given the receiver's interim beliefs, the posterior beliefs of the receiver and of the sender following a signal realisation s form a bijection that is independent of the signal realisation s and of the signalling strategy π (aside from the effect of the equilibrium choice of signalling strategy on the receiver's interim belief). This is shown in Alonso & Câmara (2016), in the context of heterogeneous priors, and naturally extends to the case of heterogeneous interim beliefs.

In particular, given interim beliefs μ_i of the sender and μ_r of the receiver, any posterior belief μ of the receiver corresponds to a posterior belief of the sender $m_s(\mu, \mu_r, \mu_i)$ defined as:

$$m_s(\mu, \mu_r, \mu_i) = \frac{\mu\mu_r(1-\mu_i)}{\mu\mu_r(1-\mu_i) + (1-\mu)(1-\mu_r)\mu_i}$$

2.3 Conditions for collusion to reveal more information than competition

The first result provides some sufficient conditions on the payoffs of the senders and on the private information structure of sender I, such that the least informative equilibrium in competition is less informative than the least informative equilibrium in collusion. In particular, I focus on one particular type of equilibrium structure such that this situation can arise. I first define formally the equilibrium and then summarise the conditions that payoffs and beliefs need to satisfy for this equilibrium to exist.

The equilibrium considered takes the following form. In competition, both types of Sender I (the informed sender) pool on an experiment, denoted π^c ('c' for 'competition'), that induces a distribution over posteriors τ^c given interim beliefs μ_0. Sender N also selects the experiment π^c. Under collusion, both types of the coalition pool on an experiment that induces a distribution over posteriors τ^m ('m' for 'merged').

I am interested in a situation in which (1) τ^m is Blackwell more informative than τ^c ($\tau^m \succsim \tau^c$), and (2) τ^m is the least informative equilibrium of the collusive game. As a result, the least informative equilibrium in collusion is more informative than the least informative equilibrium in competition. I define this situation as a *collusion-preferred* pair of equilibria.

Definition 3. *A pair of equilibrium strategies of the senders together with a belief function of the receiver* $((\boldsymbol{\pi^c}, \mu^c(\pi)), (\pi^m, \mu^m(\pi)))$ *is a* ***collusion-preferred pair of equilibria*** *if:*

1. $(\boldsymbol{\pi^c}, \mu^c(\pi))$ *is a symmetric pooling weak perfect Bayesian equilibrium of the competitive game that survives the intuitive criterion.*

2. $(\pi^m, \mu^m(\pi))$ *is a pooling weak perfect Bayesian equilibrium of the collusive game that survives the intuitive criterion.*

3. *If* $(\pi^{m'}, \mu^{m'}(\pi))$ *is another weak perfect Bayesian equilibrium of the collusive game that survives the intuitive criterion, then* $\langle \pi^{m'} \rangle \succsim \langle \pi^m \rangle$

4. $\langle \pi^m \rangle = \tau^m \succsim \tau^c = \langle \boldsymbol{\pi^c} \rangle$.

Proposition 7 summarises some sufficient conditions that the senders' payoffs should satisfy, given some private information structure for the existence of a collusion-preferred pair of equilibria. If a payoff structure satisfying these conditions exists, competition actually reduces rather than increases the amount of information produced.

Proposition 7. *Consider a pair of senders' payoff functions* $(v_I(a), v_N(a))$*, and a pair of experiments* (π^c, π^m) *such that* $\langle \pi^m \rangle \succsim \langle \pi^c \rangle$*. There exists a collusion-preferred pair of equilibria if the following conditions are satisfied:*

1. $(\boldsymbol{\pi^c}, \mu^c(\pi))$ *is a symmetric pooling weak perfect Bayesian equilibrium of the competitive game:*

   (a) *For any* π',

$$\mathbb{E}_{\mu_0}\left[V_N((\mu_s)_{(s\in\boldsymbol{\pi^c})}; \mu_0)\right] > \mathbb{E}_{\mu_0}\left[V_N((\mu_s)_{(s\in\pi'\vee\pi^c)}; \mu_0)\right] \tag{2.2}$$

   (b) *For any* $\mu_\theta \in \{\mu_{\theta_L}, \mu_{\theta_R}\}$ *and for all* π'

$$\mathbb{E}_{\mu_\theta}\left[V_I((\mu_s)_{(s\in\boldsymbol{\pi^c})}; \mu_0)\right] > \mathbb{E}_{\mu_\theta}\left[V_I((\mu_s)_{(s\in\pi'\vee\pi^c)}; \tilde{\mu}(\pi'))\right] \quad \text{for some } \tilde{\mu}(\pi') \text{ satisfying condition 1(c)} \tag{2.3}$$

   (c) $\tilde{\mu}(\pi')$ *should satisfy*

$$\begin{aligned}
&\text{Either } \tilde{\mu}(\pi') = \mu_\theta \in \{\mu_{\theta_L}, \mu_{\theta_R}\} \text{ and, both} \\
&\qquad \begin{cases} \exists\ \mu \ \text{ s.t. } \ \mathbb{E}_{\mu_\theta}\left[V_I((\mu_s)_{(s\in\pi'\vee\pi^c)}; \mu)\right] > \mathbb{E}_{\mu_\theta}\left[V_I((\mu_s)_{(s\in\boldsymbol{\pi^c})}; \mu_0)\right] \\ \text{and } \forall\ \mu, \ \ \mathbb{E}_{\mu_{\theta'}}\left[V_I((\mu_s)_{(s\in\pi'\vee\pi^c)}; \mu)\right] < \mathbb{E}_{\mu_{\theta'}}\left[V_I((\mu_s)_{(s\in\boldsymbol{\pi^c})}; \mu_0)\right] \end{cases} \\
&\quad \text{Or, } \tilde{\mu}(\pi') = \mu_0 \text{ and either} \\
&\qquad \forall\ \mu_\theta,\ \forall\ \mu,\ \ \mathbb{E}_{\mu_\theta}\left[V_I((\mu_s)_{(s\in\pi'\vee\pi^c)}; \mu)\right] < \mathbb{E}_{\mu_\theta}\left[V_I((\mu_s)_{(s\in\boldsymbol{\pi^c})}; \mu_0)\right] \\
&\quad \text{Or, } \exists\ \mu(\theta),\ \text{ s.t. } \forall \mu_\theta, \\
&\qquad\qquad \mathbb{E}_{\mu_\theta}\left[V_I((\mu_s)_{(s\in\pi'\vee\pi^c)}; \mu(\theta))\right] > \mathbb{E}_{\mu_\theta}\left[V_I((\mu_s)_{(s\in\boldsymbol{\pi^c})}; \mu_0)\right]
\end{aligned} \tag{2.4}$$

2. $(\pi^m, \mu^m(\pi))$ *is a pooling weak perfect Bayesian equilibrium of the collusive game*

   (a) *For any* $\mu_\theta \in \{\mu_{\theta_L}, \mu_{\theta_R}\}$ *and for all* π'

$$\mathbb{E}_{\mu_\theta}\left[V_m((\mu_s)_{(s\in\pi^m)}; \mu_0)\right] > \mathbb{E}_{\mu_\theta}\left[V_m((\mu_s)_{(s\in\pi')}; \tilde{\mu}(\pi'))\right] \quad \text{for some } \tilde{\mu}(\pi') \text{ satisfying condition 2(b)} \tag{2.5}$$

(b) $\tilde{\mu}(\pi')$ *should satisfy*

$$\begin{aligned}
&\textit{Either } \tilde{\mu}(\pi') = \mu_\theta \in \{\mu_{\theta_L}, \mu_{\theta_R}\} \textit{ and, both}\\
&\qquad\begin{cases} \exists\ \mu \ \ s.t.\ \ \mathbb{E}_{\mu_\theta}\left[V_m((\mu_s)_{(s\in\pi')};\mu)\right] > \mathbb{E}_{\mu_\theta}\left[V_m((\mu_s)_{(s\in\pi^m)};\mu_0)\right] \\ \textit{and } \forall\ \mu,\ \ \mathbb{E}_{\mu_{\theta'}}\left[V_m((\mu_s)_{(s\in\pi')};\mu)\right] < \mathbb{E}_{\mu_{\theta'}}\left[V_m((\mu_s)_{(s\in\pi^m)};\mu_0)\right] \end{cases}\\
&\textit{Or, } \tilde{\mu}(\pi') = \mu_0 \textit{ and either}\\
&\qquad \forall\ \mu_\theta,\ \forall\ \mu,\ \ \mathbb{E}_{\mu_\theta}\left[V_m((\mu_s)_{(s\in\pi')};\mu)\right] < \mathbb{E}_{\mu_\theta}\left[V_m((\mu_s)_{(s\in\pi^m)};\mu_0)\right]\\
&\textit{Or, } \exists\ \mu(\theta),\ \ s.t.\ \forall\mu_\theta,\\
&\qquad\qquad \mathbb{E}_{\mu_\theta}\left[V_m((\mu_s)_{(s\in\pi')};\mu(\theta))\right] > \mathbb{E}_{\mu_\theta}\left[V_m((\mu_s)_{(s\in\pi^m)};\mu_0)\right]
\end{aligned} \tag{2.6}$$

3. $(\pi^m, \mu^m(\pi))$ *is the least informative weak perfect Bayesian pooling equilibrium of the collusive game: for any* π' *such that* $\langle\pi^m\rangle \succsim \langle\pi'\rangle$,

   (a) $\exists\ \mu_\theta$ *and* $\exists\ \pi''$ *such that*

$$\mathbb{E}_{\mu_\theta}\left[V_m((\mu_s)_{(s\in\pi'')};\tilde{\mu}(\pi''))\right] > \mathbb{E}_{\mu_\theta}\left[V_m((\mu_s)_{(s\in\pi')};\mu_0)\right] \tag{2.7}$$

   (b) *Condition 3(a) holds for any* $\tilde{\mu}(\pi'')$ *such that*

$$\begin{aligned}
&\textit{Either } \tilde{\mu}(\pi'') = \mu_\theta \in \{\mu_{\theta_L}, \mu_{\theta_R}\} \textit{ and, both}\\
&\qquad\begin{cases} \exists\ \mu \ \ s.t.\ \ \mathbb{E}_{\mu_\theta}\left[V_m((\mu_s)_{(s\in\pi'')};\mu)\right] > \mathbb{E}_{\mu_\theta}\left[V_m((\mu_s)_{(s\in\pi')};\mu_0)\right] \\ \textit{and } \forall\ \mu,\ \ \mathbb{E}_{\mu_{\theta'}}\left[V_m((\mu_s)_{(s\in\pi'')};\mu)\right] < \mathbb{E}_{\mu_{\theta'}}\left[V_m((\mu_s)_{(s\in\pi')};\mu_0)\right] \end{cases}\\
&\textit{Or, } \tilde{\mu}(\pi'') = \mu_0 \textit{ and either}\\
&\qquad \forall\ \mu_\theta,\ \forall\ \mu,\ \ \mathbb{E}_{\mu_\theta}\left[V_m((\mu_s)_{(s\in\pi'')};\mu)\right] < \mathbb{E}_{\mu_\theta}\left[V_m((\mu_s)_{(s\in\pi')};\mu_0)\right]\\
&\textit{Or, } \exists\ \mu(\theta),\ \ s.t.\ \forall\mu_\theta,\\
&\qquad\qquad \mathbb{E}_{\mu_\theta}\left[V_m((\mu_s)_{(s\in\pi'')};\mu(\theta))\right] > \mathbb{E}_{\mu_\theta}\left[V_m((\mu_s)_{(s\in\pi')};\mu_0)\right]
\end{aligned} \tag{2.8}$$

4. *Any separating equilibrium in collusion, if it exists, reveals more information than* π^m.

The first set of conditions (1(a), 1(b) and 1(c)) ensures the existence of a pooling equilibrium in competition in which the strategy profile $\boldsymbol{\pi^c}$ is played. The second set of conditions (2(a) and 2(b)) ensures that there is a pooling equilibrium in the collusive game in which the experiment $\boldsymbol{\pi^m}$ is played. The third set of conditions rules out the existence of a pooling equilibrium in collusion that is less informative than the experiments played in competition. Finally, the last condition allows us to focus on pooling equilibria as any separating equilibrium is more informative than π^m.

In particular, condition 1(a) (inequality 2.2) guarantees that Sender N does not deviate to any other experiments given that Sender I plays π^c. Since Sender N has no information in competition, the out-of-equilibrium beliefs of the receiver are always equal to the prior μ_0 following any deviation by Sender N.

Condition 1(b) (inequality 2.3) ensures that no type of Sender I wants to deviate from the equilibrium experiment, if that deviation were to induce out-of-equilibrium belief $\tilde{\mu}(\pi')$. In addition, this out-of-equilibrium belief needs to satisfy the intuitive criterion. This is guaranteed by condition 1(c) (inequalities 2.4). In particular, the receiver should put weight only on type μ_θ, following a deviation to π', if there is a belief such that this deviation gives a higher payoff than the equilibrium payoff for type μ_θ, but there is no such belief for type $\mu_{\theta'}$. Otherwise, if either both types could potentially benefit from the deviation, or if no type would, then the receiver's belief should equal the prior.

Similarly, condition 2(a) (inequality 2.5) ensures that the coalition does not have incentives to deviate to an alternative experiment π' if that deviation were to induce out-of-equilibrium belief $\tilde{\mu}(\pi')$. Condition 2(b) (inequalities 2.6) defines the set of out-of-equilibrium beliefs $\tilde{\mu}(\pi')$ that satisfy the intuitive criterion.

Condition 3(a) (inequality 2.7) guarantees that π^m is the least informative pooling equilibrium in collusion. For any other experiment π', one type would find it profitable to deviate to another experiment π'' for any out-of-equilibrium belief $\tilde{\mu}(\pi'')$ that satisfies the intuitive criterion.

Finally, condition 4 can be satisfied whenever the informed sender's private information is sufficiently precise that the signalling effect in itself would reveal more information than the equilibrium choice of experiment.

As a result, when all these conditions are satisfied, the least informative equilibrium of the competitive game is less informative than any equilibrium of the collusive game.

2.3.1 Alternative combinations of equilibria

Proposition 7 offers some *sufficient* conditions for the existence of *some* combination of equilibria in which the least informative equilibrium under collusion is more informative than the least informative equilibrium in competition. This is not the only possible combination of equilibria that would lead to this situation, however. In particular, it is possible that the conditions guaranteeing this pair of equilibria are not satisfied, but there exist other equilibria that lead to more information under collusion than competition, even when restricting attention to the least informative equilibria.

I focus on this situation because it captures the most interesting aspect of persuasion with private information. In particular, it isolates two effects: one due to the uninformed sender learning private information in collusion, and one due to the receiver having different off-equilibrium beliefs under collusion than under competition. In addition, it excludes a third possible effect: increasing informativeness through the revelation of the sender's type. Indeed, because Proposition 7 focuses on pooling equilibria, such learning does not occur in equilibrium and the increase in informativeness comes exclusively from the equilibrium choice of experiments.

Since I focus on least informative equilibria, separating equilibria can be easily ruled out from the comparison provided that the informed sender's information is sufficiently precise. In particular, if that information in itself is more informative than the experiment on which the senders pool (both in collusion and competition), then a separating equilibrium can never be the least informative equilibrium.

Finally, in competition, the same outcome distribution of posterior beliefs can be generated from different individual choice of experiments. To keep the problem simple, I focus on a symmetric equilibrium in competition: both senders choose the same experiment.

2.4 Sufficient conditions for selected payoff functions

To understand the dynamics involved in a collusion-preferred pair of equilibria and to show that such a situation can exist, I restrict attention to payoffs functions of a particular form. This simpler problem highlights the key differences that arise between competition and collusion when senders are privately informed while limiting the set of actions of the senders and the set of possible equilibria.

I consider senders with discontinuous payoff functions who have opposite preferences on some range of receiver's actions: when a sender prefers the receiver to have posterior beliefs closer to 1, the other sender prefers the receiver to have posterior beliefs closer to 0 (but only up to a certain limit).

Assumption 1. *The preferences of the senders satisfy the following assumption:*

$$v_I(\mu) = \begin{cases} \frac{w_1^I}{\mu_1}\mu \text{ on } [0, \mu_1] \\ w_2^I \text{ on } (\mu_1, \mu_2] \\ w_3^I \text{ on } (\mu_2, 1] \end{cases} \qquad v_N(\mu) = \begin{cases} w_1^N \text{ on } [0, \mu_1) \\ w_2^N \text{ on } [\mu_1, \mu_2) \\ w_3^N \text{ on } [\mu_2, 1] \end{cases} \tag{2.9}$$

Where $0 < \mu_1 < \mu_0 < \mu_2 < 1$, *and* $w_1^i < w_2^i < w_3^i$ *and* $w_3^j < w_2^j < w_1^j$ *for* $i, j \in \{I, N\}$.

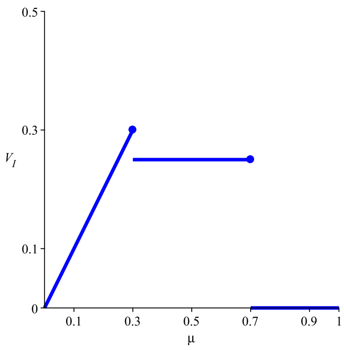
Notice that these assumptions imply that the two senders are competing, in the sense that one sender prefers the receiver to have posterior beliefs closer to 1 whereas the other sender prefers the receiver to have posterior beliefs closer to 0. However, the senders have partially aligned interests at the points μ_1 and μ_2 since they both get relatively high utility at these boundaries.

Under collusion, the coalition's utility is a convex combination of the individual senders' utilities with a weight α on the preferences of Sender I. That is,

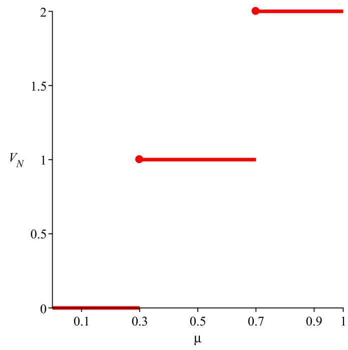
$$v^c(\mu) = \begin{cases} w_1^c = \alpha \cdot \left(\frac{w_1^I}{\mu_1}\mu\right) + (1-\alpha) \cdot w_1^N & \text{on } [0, \mu_1) \\ w^c(\mu_1) = \alpha \cdot w_1^I + (1-\alpha) \cdot w_2^N & \text{at } \mu_1 \\ w_2^c = \alpha \cdot w_2^I + (1-\alpha) \cdot w_2^N & \text{on } (\mu_1, \mu_2) \\ w^c(\mu_2) = \alpha \cdot w_2^I + (1-\alpha) \cdot w_3^N & \text{at } \mu_2 \\ w_3^c = \alpha \cdot w_3^I + (1-\alpha) \cdot w_3^N & \text{on } [\mu_2, 1] \end{cases} \tag{2.10}$$

These utility functions are illustrated in figure 2.2. Note that the points of discontinuity at μ_1 and μ_2 in the coalition's utility functions are due to the partial overlap of preferences by the senders at these points. This feature of the payoff functions are needed to ensure existence of a non-fully revealing equilibrium in competition, but it does not drive the desire for a more informative distribution in collusion. Intuitively, these points correspond to regions where the senders are in relative agreement: they would prefer to push the posterior beliefs further in one direction or the other, but they agree that this compromise is better than the next worse option. The results would also work with a larger overlap around these beliefs.

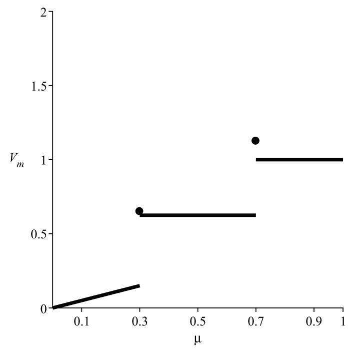
I also impose restrictions on the private information of the informed sender. As described in section 2.3.1, separating equilibria can be excluded from the analysis of the least informative equilibria if the signalling itself always reveal more information than the experiments chosen in pooling equilibria.



(a) Sender I



(b) Sender N



(c) Coalition ($\alpha = 0.4$)

Figure 2.2: Payoff functions

I assume that the private information of the informed sender is sufficiently informative that the interim belief of this sender is either below μ_1 or above μ_2. In particular, the signal realisation θ_R fully reveals the state, whereas realisation θ_L induces an interim belief between 0 and μ_1:

Assumption 2. *$p(\theta|\omega)$ is such that:*

$$\begin{aligned} p(\theta_R|L) &= 0 \\ p(\theta_L|R) &< \frac{1-\mu_0}{\mu_0}\frac{\mu_1}{1-\mu_1} \end{aligned} \tag{2.11}$$

Assumption 2 implies that $\mu_{\theta_L} \in (0, \mu_1)$ and $\mu_{\theta_R} = 1$.

2.4.1 Conditions on payoffs for the existence of a collusion-preferred pair of equilibria

Given these restrictions, the next step is to find conditions on the senders payoffs and private information such that all conditions in Proposition 7 are satisfied. In particular, one way to achieve this is to look for value of these parameters such that the equilibrium experiments induce distributions $\langle \pi^m \rangle = \tau^m$ and $\langle \boldsymbol{\pi^c} \rangle = \tau^c$ that satisfy

$$\tau^c = [\mu_1 \text{ w.p. } \tau_1^c;\ \mu_2 \text{ w.p. } \tau_2^c]$$
$$\tau^m = [0 \text{ w.p. } \tau_1^m;\ \mu_2 \text{ w.p. } \tau_2^m]$$

Under these conditions, $\tau^m \succsim \tau^c$, since one signal realisation induces the same belief (μ_2) under both distributions, while the other realisation fully reveals the state in collusion (induces belief 0), but not in competition (induces belief $\mu_1 > 0$).

The equilibria inducing these distributions take the following form. In collusion, both types of the coalition choose an experiment π^m such that

$$\mathbb{P}_{\pi^m}(s_1|R) = 0 \text{ and } \mathbb{P}_{\pi^m}(s_1|L) = \frac{\mu_2 - \mu_0}{\mu_2(1-\mu_0)}$$

In competition, there exists a pair of symmetric experiments (π^c, π^c) that results in two possible signal realisations s_1 and s_2 such that:

$$\mathbb{P}_{(\pi^c \vee \pi^c)}(s_1|R) = \frac{\mu_1(\mu_2 - \mu_0)}{\mu_0(\mu_2 - \mu_1)} \text{ and } \mathbb{P}_{(\pi^c \vee \pi^c)}(s_1|L) = \frac{(1-\mu_1)(\mu_2 - \mu_0)}{(1-\mu_0)(\mu_2 - \mu_1)}$$

It is possible for two symmetric experiments to have only two signal realisations because the rich signal space considered here allows senders to perfectly correlate their signal realisations to the other sender's realisation. The distribution above can therefore be achieved if the two senders' expriments are perfectly correlated and issue a signal realisation s_1 with probability $\frac{\mu_1(\mu_2-\mu_0)}{\mu_0(\mu_2-\mu_1)}$ when the state is R and probability $\frac{(1-\mu_1)(\mu_2-\mu_0)}{(1-\mu_0)(\mu_2-\mu_1)}$ when the state is L. Because the two signal realisations are perfectly correlated, the receiver does not gain additional information from observing the two realisations and a realisation s therefore induces the same belief whether it comes from a single experiment π^c or from a pair of such experiments. In addition, if the two signal realisations are perfectly correlated, there is a zero probability that the receiver observes realisation s_1 from one sender's experiment and s_2 from another, so the support of the combined signals is also binary. The construction of these signals is provided in appendix.

In both games, the receiver's interim beliefs are equal to the prior after observing an equilibrium experiment, and are equal to any beliefs in the set of beliefs consistent with the intuitive criterion following a deviation.

Necessary condition on senders' payoffs

For $\boldsymbol{\pi^c}$ and π^m to be equilibria, each individual sender must prefer the less informative distribution (τ^c) in competition but the coalition of the two senders must prefer the more informative distribution (τ^m), at least for some type of the informed sender.

In particular, since the signal θ_R is fully revealing, the informed sender cannot get a higher payoff when the receiver knows that the state is R (i.e. $\mu = 1$). Otherwise, type θ_R would always deviate to a fully revealing experiment (or any experiment that signals his type), and induce $\mu = 1$ (given $\mu_{\theta_R} = 1$).

It should also be the case that the coalition does not get a higher payoff when the receiver is certain that the state is R. Otherwise type θ_R of the coalition would deviate to a fully revealing experiment.

Lemma 3. *If $\boldsymbol{\pi^c}$ is an equilibrium of the competitive game, and π^m an equilibrium of the collusive game, then*

1. $w_1^I > w_2^I > w_3^I$ *for any i, j such that $i > j$.*

2. $\max\{w_2^c, w_1^c, w^c(\mu_1), w^c(\mu_2)\} > w_3^c$

A direct corollary of Lemma 3 and assumption 1 is that Sender N has increasing preferences: $w_1^N < w_2^N < w_3^N$.

Sufficient conditions on senders' payoffs

I now derive some sufficient conditions for the existence of a collusion-preferred equilibrium.

Proposition 8. *Suppose that preferences of the senders satisfy the conditions in Lemma 3, and normalise the senders' payoffs so that $w_1^N = w_3^I = 0$, then a collusion-preferred equilibrium exists if $w_1^c < w_2^c < w_3^c$ and,*

$$\frac{\mu_1}{\mu_2} \leq \frac{w_2^N}{w_3^N} \tag{2.12}$$

$$\frac{(1-\mu_2)(\mu_0(1-\mu_1) - \mu_{\theta_L}(\mu_0 - \mu_1))}{(1-\mu_1)(\mu_0(1-\mu_2) + \mu_{\theta_L}(\mu_2 - \mu_0))} \leq \frac{w_2^I}{w_1^I} \tag{2.13}$$

$$\frac{\mu_{\theta_L}\mu_2(1-\mu_0)}{\mu_1(\mu_{\theta_L}\mu_2(1-\mu_0) + \mu_0(1-\mu_{\theta_L})(1-\mu_2))} \leq \frac{w^c(\mu_2)}{w^c(\mu_1)} \tag{2.14}$$

$$\frac{\mu_1(\mu_2-\mu_0)}{\mu_0(\mu_2-\mu_1)} > \frac{w^c(\mu_2) - w_3^c}{w^c(\mu_2) - w^c(\mu_1)} \tag{2.15}$$

Competition. Condition 2.12 ensures no deviation by Sender N from π^c in competition. By Lemma 3, we know that Sender N's preferences must be increasing. Therefore, Sender N would only want to deviate to a strategy that increases the weight on μ_2. This can only be achieved by 'splitting' belief μ_1 and condition 2.12 ensures that this is not profitable. In particular, condition 2.12 implies that μ_1 and μ_2 are on the concave closure of Sender N's utility so any other distribution of posterior beliefs would yield a lower expected utility.

Condition 2.13 ensures that type θ_L of Sender I does not deviate to induce beliefs $(\mu_1, 1)$ with an experiment that induces out-of-equilibrium beliefs μ_0. The concavification approach cannot be used directly when the sender is privately informed since the sender's expected utility is evaluated based on his belief μ_θ rather than the prior μ_0 (as was the case for Sender N). However, the same intuition can be used to rule out any distributions that puts weight on any belief other than μ_1, μ_2 or 1. Condition 2.13 then ensures that type θ_L prefers to induce (μ_1, μ_2) than $(\mu_1, 1)$, and is sufficient for type θ_L not to deviate to any other experiment.

A similar condition ensures that type θ_R of Sender I does not deviate to induce beliefs $(\mu_1, 1)$ with an experiment that induces out-of-equilibrium beliefs μ_0. Since type θ_L is more confident that the state is favourable than type θ_R, the requirement is less demanding for type θ_R. Therefore, condition 2.13 is also sufficient to ensure that type θ_R does not deviate.

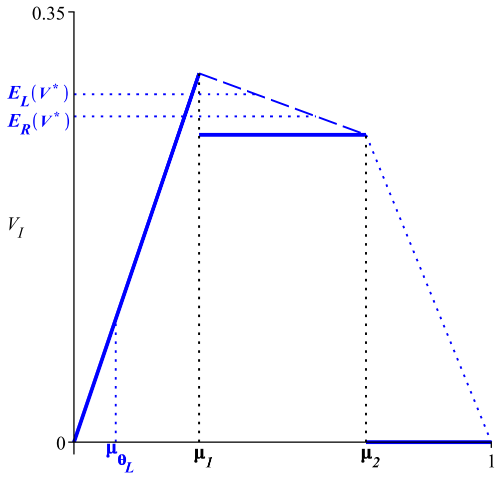
If some deviation induces out-of-equilibrium belief μ_{θ_R}, then the receiver's posterior belief is $\mu = 1$ for any experiment. Therefore, Sender I's expected utility is w_3^I and since $w_3^I < w_2^I < w_1^I$, Sender I does not want to deviate to such an experiment.

If some deviation induces out-of-equilibrium belief μ_{θ_L}, condition 2.13 ensures that type θ_L does not want to deviate to that experiment. In particular, in that case, the set of possible deviations for type θ_L is restricted by (1) the information revealed by Sender N playing π^c, and (2) the revelation that the informed sender has observed θ_L. This implies that any deviation must induce some beliefs in the set $[0, m(\mu_1, \mu_0, \mu_{\theta_L})] \cup [m(\mu_2, \mu_0, \mu_{\theta_L}), 1]$. The conditions for no deviation to be profitable can be reduced to two conditions by noting that any deviations inducing a distribution over more than two posterior beliefs is dominated by either inducing $(m(\mu_1, \mu_0, \mu_{\theta_L}), \mu_2)$ or $(m(\mu_1, \mu_0, \mu_{\theta_L}), 1)$. Finally, these two conditions can be reduced to a unique one (condition 2.13) as one implies the other.

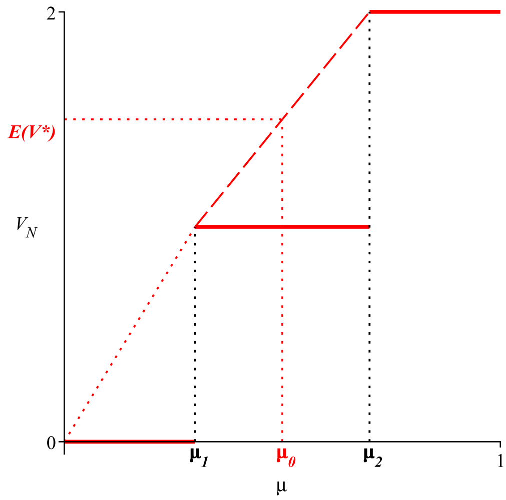
Note that there is no need to check whether type θ_R wants to deviate to an experiment, if that deviation induces out-of-equilibrium belief μ_{θ_L}, since if he does, then the receiver's set of beliefs satisfying the intuitive criterion following that deviation includes μ_{θ_R} and the deviation can be ruled out by associating belief μ_{θ_R} with that deviation.

Therefore, taken together, conditions 2.12 and 2.13 guarantee that the strategy profile (π^c, π^c) is an equilibrium of the competitive game.

Figure 2.3 illustrates the concave closure of the senders' expected utility functions and their expected payoffs in the competitive equilibrium. Sender N has no information, so his expected utility in this pooling equilibrium (denoted $E(V^*)$ in the graph) is evaluated with the same beliefs as those of the receivers (μ_0). Sender I has different beliefs than the receiver in equilibrium. However, the beliefs induced are on the concave closure of the sender's expected utility given the receiver's interim belief μ_0. The only type of the sender who might want to deviate is type θ_L, but as illustrated, his equilibrium utility (denoted $E_L(V^*)$) is greater than the best deviation that reveals his type (and which must therefore



(a) Sender I



(b) Sender N

Figure 2.3: Competitive equilibrium

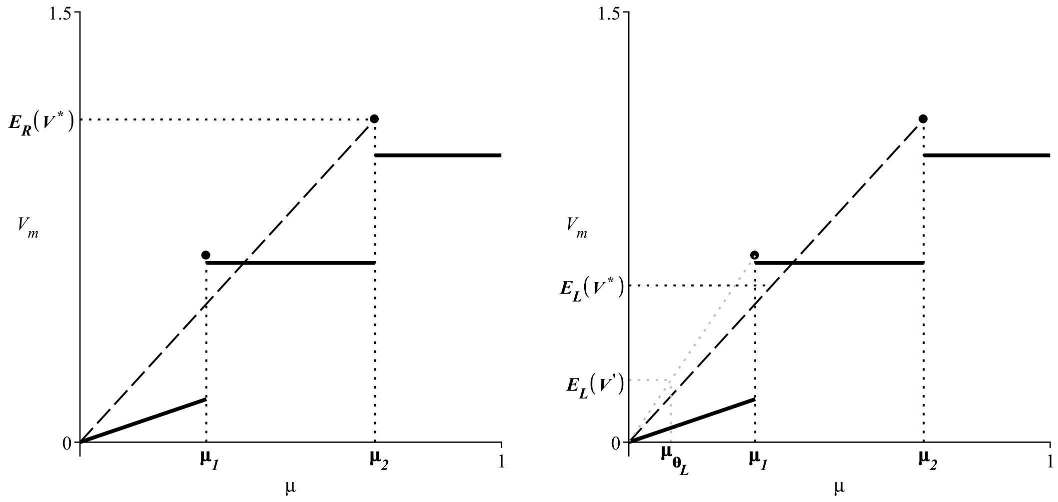
induce some belief to the left of μ_{θ_L}).

Collusion. In collusion, the two types of the coalition pool on an experiment inducing posterior beliefs $(0, \mu_2)$. First note that in this case, we can focus exclusively on deviations from type θ_L which induce out-of-equilibrium beliefs μ_{θ_L}. Indeed, under the equilibrium strategy, type θ_R receives the highest payoff, $w^c(\mu_2)$, with probability 1. Therefore, no deviation can be profitable for type θ_R, and the set of out-of-equilibrium beliefs that survive the intuitive criterion only includes μ_{θ_L}.

Condition 2.14 ensures that type θ_L of the coalition does not want to deviate from π^m in collusion. In particular, since any such deviations induces out-of-equilibrium belief μ_{θ_L}, the set of feasible beliefs that can be induced by these deviations is $[0, \mu_{\theta_L}] \cup [\mu_{\theta_L}, 1]$. Following a standard concavification argument, the best possible deviations given this set induce either (m_1, μ_1) or (m_1, μ_2) for some $m_1 \in [0, \mu_{\theta_L}]$. Condition 2.14 is sufficient for both of these deviations to be dominated by the equilibrium payoff.

Finally, pooling on an experiment that induces a less informative distribution of posterior beliefs than $\langle \pi^c \rangle$ is not an equilibrium in collusion when condition 2.15 is satisfied. In particular, if the inequality holds, type θ_R would always prefer to deviate from an equilibrium inducing beliefs (μ_1, μ_2) to a fully revealing experiment and receive payoff w_3^c with certainty. Any other experiment inducing a less informative distribution of beliefs is not an equilibrium as some type of the coalition would deviate to induce beliefs $(0, \mu_1)$ for any out-of-equilibrium beliefs that satisfy the intuitive criterion.

Figure 2.4 illustrates the concave closure of the coalition's expected utility functions



(a) Expected payoff of type θ_R (b) Expected payoff of type θ_L

Figure 2.4: Collusive equilibrium

and its expected payoff in the collusive equilibrium. Type θ_R gets the highest possible payoffs since he expects belief μ_2 to be induced with probability 1 and to receive w_3^c (denoted $E_R(V*)$ in the graph). Type θ_L receives expected utility $\mathbb{E}_{\mu_{\theta_L}}[V_C(0, \mu_2; \mu_0)]$ (denoted $E_L(V^*)$ in the graph), which is higher than the utility from the best deviation which reveals his type (denoted $E_L(V^{'})$).

2.4.2 Existence of more informative collusive equilibrium

The next proposition shows that when the information structure satisfies assumption 2 then the set of payoffs that satisfy all the conditions in Proposition 7 is non-empty.

Proposition 9. *There exists a private information structure that satisfies assumption 2 and preferences of the senders satisfying assumption 1 such that a collusion-preferred pair of equilibria exists.*

This statement is proved using a numerical example. Consider two senders with the following preferences:

$$v_I(\mu) = \begin{cases} \mu \text{ if } \mu \in [0, 0.3] \\ 0.25 \text{ if } \mu \in (0.3, 0.7] \\ 0 \text{ if } \mu \in (0.7, 1] \end{cases} \qquad v_N(\mu) = \begin{cases} 0 \text{ if } \mu \in [0, 0.3) \\ 1 \text{ if } \mu \in [0.3, 0.7) \\ 2 \text{ if } \mu \in [0.7, 1] \end{cases} \tag{2.16}$$

So that $\mu_1 = 0.3$ and $\mu_2 = 0.7$. In addition, suppose that the preferences of both senders are weighted equally in the coalition: $\alpha = \frac{1}{2}$, so the coalition's utility function is $v_m(\mu) = \frac{v_I(\mu)+v_N(\mu)}{2}$. These payoffs are those depicted in figure 2.2 and satisfy assumption 1.

Sender I receives some information privately before designing the experiment: a binary signal $\theta \in \{\theta_L, \theta_R\}$ distributed according to some conditional distribution $p(\theta|\omega)$ and satisfying assumption 2. In particular, suppose that

$$p(\theta_L|R) = \frac{1}{9} \text{ and } p(\theta_L|L) = 1, \tag{2.17}$$

so that $\mu_{\theta_L} = \frac{1}{10}$ and $\mu_{\theta_R} = 1$.

These parameter values satisfy the conditions in Proposition 8, so the least informative collusive equilibrium is more informative than the least informative competitive equilibrium.

In particular, neither type of Sender I want to deviate to an experiment inducing a more informative distributions in competition. Sender N also does not want to deviate to an experiment inducing a more informative distribution in the absence of information.

In collusion, Sender N learns about the private information of Sender I and the coalition now prefers the more informative distribution of posterior beliefs when they observe private signal θ_R. Type θ_L would prefer to deviate to an experiment inducing (μ_1, μ_2) given interim μ_0, but since type θ_R does not, the intuitive criterion only allows the receiver to put weight on type θ_L following that deviation. Type θ_L could deviate by inducing a less informative distribution of posterior beliefs given interim μ_{θ_L} but this deviation is unprofitable for the coalition. As a result, pooling on an experiment inducing the more informative distribution is an equilibrium.

Finally, if the coalition were to pool on an experiment inducing the less informative distribution over posterior beliefs τ^c, then type θ_R would find it profitable to deviate to a fully-revealing experiment, as this type expects a fully-revealing experiment to guarantee a payoff of $w_3^c = 1$, independently of the out-of-equilibrium beliefs of the receiver (these beliefs no longer matter with a fully revealing experiment). Type θ_R therefore deviates and pooling on an experiment inducing the less informative distribution of posterior beliefs is not an equilibrium. This arises in collusion but not in competition because the coalition's preferences put more weight on the preferences of the uninformed sender, who would not deviate in this way in competition because he does not have private information.

2.5 Comparative statics

2.5.1 Role of private information

Since both senders have access to the private information of the informed sender when they collude, the perception of the outcome for the uninformed sender changes, but also the out-of-equilibrium beliefs of the receiver following a deviation. The latter effect occurs as the equilibrium expected utility and the expected utility from deviating are different in the coalition than for the informed sender alone. However, only certain combinations of payoffs and private information structure allow all the conditions in Proposition 7 to be satisfied at the same time. In particular, the senders' private information needs to be sufficiently informative, but not too informative for these conditions to be satisfied.

Proposition 10. *If neither sender has any private information, i.e. if* $\frac{p(\theta_R|R)}{p(\theta_R|L)} = 1$*, then a collusion-preferred pair of equilibria does not exist.*

This result follows directly from Gentzkow & Kamenica (2017*b*): when the information environment is Blackwell connected (as is the case here), the least informative equilibrium in competition is always more informative than the least informative equilibrium in collusion. It is also easy to verify directly that the collusion-preferred pair of equilibria cannot exist in this case, because there is a contradiction between conditions 2.2, 2.3 and 2.5. Conditions 2.2 and 2.3 require both senders to prefer the equilibrium distribution τ^c to the more informative distribution τ^m. But if this holds, the coalition would also prefer the equilibrium distribution τ^c thus contradicting condition 2.5.

More generally, it needs to be the case that at least one of the senders, when provided with some private information, can gain from generating a more informative distribution than the competitive equilibrium distribution. In addition, it should also be true that unilaterally deviating to that more informative distribution is not profitable to the uninformed sender. Otherwise, he could deviate to an experiment inducing it, without changing the off-equilibrium beliefs of the receiver, since he has no private information to reveal through his action.

In addition, the amount of information that sustains a collusion-preferred pair of equilibria is bounded: if both senders observed a fully revealing private signal, and if preferences satisfy assumption 1, a collusion-preferred pair of equilibria cannot exist.

Proposition 11. *If both senders privately observe fully revealing signals, then there does not exist a set of payoffs satisfying assumption 1 such that a collusion-preferred pair of equilibria exists.*

With fully revealing private signals, the interim beliefs of the senders are either 0 or 1. Given lemma 3, we know that preferences need to be increasing for one sender and decreasing for the other if a collusion-preferred equilibrium is to exist.

Suppose that Sender I's preferences are decreasing while Sender N's preferences are increasing, and that both senders are perfectly informed. Consider a competitive pooling equilibrium experiment that induces posterior beliefs (m_1, m_2) with $m_2 < 1$, $m_1 < \mu_1$. Then type θ_L of Sender I would deviate to an experiment inducing $(m_1, 1)$ as this guarantees that m_1 will be induced with probability 1 and yields the highest payoff. Similarly, if the competitive equilibrium experiment induces posterior beliefs (m_1, m_2) with $m_1 > 0$, $m_2 > \mu_2$, type θ_R of Sender N would deviate to an experiment inducing $(0, m_2)$ to guarantee m_2 and yield the highest payoff. So pooling on an equilibrium revealing less information than the collusive equilibrium is not possible with full information.

In addition, any separating equilibrium in competition would fully reveal the senders' types, so the collusive equilibrium cannot be strictly more informative than the competitive equilibrium.

Private information to both senders. The assumption that only one sender observes private signals before designing an experiment can be relaxed. In the example above, the coalition always wants to reveal more information than in competition, because the preferences of the uninformed sender are such that one type of the coalition gains sufficiently from revealing that information that it wants to deviate. However, if that information was not sufficiently precise, it is still possible that this sender would prefer the less informative experiment. Therefore, the situation described in the example above could arise if the uninformed sender had access to some sufficiently imperfect signal when competing, but gained access to more precise information when colluding.[4]

2.5.2 Welfare effects

I now turn to evaluating the welfare of all players, and how the presence of private information affects it. Throughout this section, I evaluate welfare at the least informative equilibrium under both collusion and competition.

[4] In addition, it is not necessary that the senders have private information in both the competitive and the collusive situation. For instance, with one sender having private information in competition, but the coalition having no private information, it is still possible that the collusive outcome is more informative than the competitive outcome. Similarly, the more informative outcome in collusion can arise when both senders are symmetrically informed in both the competitive and collusive situations, and when the senders gain information only in collusion (for instance, if they can pool some financial resources to acquire information in collusion but not individually in competition).

If more information benefits the receiver[5], then the receiver is better-off when senders collude if the conditions such that a collusion-preferred equilibrium are satisfied.

Proposition 12. *Suppose the conditions in Proposition 8 are satisfied, then the receiver is always better-off under collusion than under competition.*

Next, I show that collusion makes the senders worse-off if it leads them to reveal more information.

Proposition 13. *Suppose the conditions in Proposition 8 are satisfied, then both senders are worse-off under collusion than under competition.*

The uninformed sender must be worse-off with the more informative distribution of posterior beliefs. If this was not the case, he could have deviated to an experiment generating that more informative distribution under competition, since his deviation has no impact on the receiver's interim beliefs.

The informed sender is also worse-off under collusion if collusion reveals more information. Suppose that, under competition, deviating to an experiment generating the more informative distribution (τ^m) induces out-of-equilibrium beliefs μ_0. If that distribution of beliefs yielded a higher payoff than the less informative distribution (τ^c), the sender would deviate to generating that distribution in competition.

If such a deviation instead induced interim μ_{θ_L}, and the informed sender preferred to induce that distribution, the deviation would be even more profitable, since the informed sender prefers the receiver to have lower posterior beliefs.

If type θ_L finds that deviation profitable, then putting weight on type θ_R only following some deviation is not consistent with the refinement used in this model. Therefore this type of the uninformed sender would always deviate to induce the more informative distribution of posteriors in competition if that distribution was more valuable to him, which contradicts the fact that the less informative distribution is an equilibrium of the competitive game.

This only leaves the possibility that type θ_R finds the more informative distribution profitable. However, type θ_R expects to get the second-highest payoff with probability 1 in collusion and a mixture of the highest and second-highest payoff in competition, so is clearly better-off in competition.

[5] For example, suppose that the receiver wants her action to match the state and has a quadratic loss utility function. This implies that the receiver's action is equal to her posterior belief. Specifically, let f a random variable measurable with respect to $\{L, R\}$, and defined by $f(R) = 1$ and $f(L) = 0$. Let $U(a, \omega) = -\mathbb{E}\left[(a - f(\omega))^2\right]$, the receiver's expected utility. This expected utility is maximised at $a = \mathbb{E}[f(\omega)] = \mu$. These preferences imply that the receiver always prefers more information.

When the conditions in Proposition 8 are satisfied but neither sender observes private signals, both senders pool on the less informative experiment both in competition and in collusion (Proposition 10). Since the senders prefer the less informative distribution (τ^c), the senders would be better-off without private information. This result is in line with existing result in the case of one informed sender (e.g. Alonso & Câmara (2018)): private information cannot make the sender better-off but can make him worse-off.

Corollary 4. *Suppose the conditions in proposition 8 are satisfied, then the senders are worse-off with private information.*

Here, both senders are worse-off with private information because the uninformed sender gains access to the information of the informed sender in collusion. In other words, the senders would gain from limiting communication in the merged organisation.

Incentives to collude and share information. Given that the senders are both worse-off, they would not agree to collude if they were given the choice. However, since the receiver is better-off under collusion, the receiver could offer transfers to the senders to induce them to collude. This would be less costly for the receiver than to offer transfers to induce competing senders to reveal more information, since it allows her to take advantage of the changes in incentives arising from senders learning each others' private information. The receiver might also have some institutional power over the organisation of the two senders. Finally, the senders might have some alternative (unmodelled) concerns that leads them to merge (such as increasing their market power).

2.6 Conclusion

This paper offers a first look at the impact of private information on the strategic choice of persuasion strategies by competing senders. In particular, it shows that access to private information by the senders may overturn a standard result about competition in persuasion: that competition tends to increase the amount of information available to the receiver.

When competing senders collaborate or collude, not only are their preferences different, but their private information also changes. This private information can induce the merged senders to deviate from any low-information equilibrium that could be sustained in competition, in order to signal that private information. As a result, the least informative equilibrium in collusion can reveal more information than the least informative equilibrium when senders compete.

There can be other forces driving a more informative equilibrium in collusion than in competition than the ones presented in this paper. For instance, collusion might force senders to separate and reveal their type, while they would prefer to pool on a common experiment and restrict the information disclosed in competition.

This paper focused on identifying sufficient conditions in one particular case, in order to analyse explicitly the mechanisms at play in this situation. This example was chosen because of its tractability, but also because it does not rely on trivial dynamics for more information to be revealed in collusion. In particular more information can be revealed with privately informed senders even when all relevant equilibria are pooling.

A full characterisation of preferences and information structures under which more information is released under collusion than competition would be an interesting avenue left for future research.

Chapter 3

Overconfidence, political accountability and politician selection

3.1 Introduction

> One of the most striking contributions to the political science of half a century of survey research has been to document how poorly ordinary citizens approximate a classical ideal of informed democratic citizenship.

This observation from Bartels (1996) summarises a large literature in political science that has shown that voters lack the ability to be well informed. Lack of political awareness can take many forms. Voters have been shown to be unaware of basic political information (Campbell et al. 1960, Delli Carpini & Keeter 1996), to vote consistently differently than fully informed agents would (Bartels 1996), to follow heuristics instead of information (Lau & Redlawsk 2001) and to react disproportionately to irrelevant information (Achen & Bartels 2004, Wolfers 2007, Leigh 2009, Healy et al. 2010) or information that is more recent (Huber et al. 2012). Several experimental studies also suggest that voters systematically misinterpret the information they receive. This includes mis-allocation of responsibilities (Hobolt et al. 2013), or systematically biased assessments of political influence (Bausch & Zeitzoff 2014, Caplan et al. 2013). A poorly informed electorate can affect policy through different channels. Voters can lack the capacity to choose politicians that will act in the public's interest, because they do not know what policies are in their own interest. They can be poorly equipped to elect competent candidates because they are unable to disentangle competence from luck when observing political outcomes. They can also be ineffective at holding policy makers accountable for choosing detrimental policies, for breaking their electoral campaign promises, or for pursuing damaging behaviour such as accepting bribery.

In this paper, I evaluate the effect of one common behavioural trait, the belief that one's information is more precise than it really is, on the welfare of voters. I show that this bias always makes voters worse at selecting good politicians under uncertainty, but can improve the incentives of politicians to take welfare improving actions. This improvement can be sufficiently large that the voter is better-off when she mis-evaluates the quality of her information.

Informed voting is important because the well functioning of democracy depends crucially on the ability of voters to scrutinise the actions or the performance of politicians. A vast empirical literature analyses the effects of voter information on political accountability including Besley & Case (1995*a*), Besley & Case (1995*b*), Besley & Burgess (2002), Besley & Case (2003), Besley (2004), Snyder & Stromberg (2010), Alt et al. (2011), and Ferraz & Finan (2011). This ability depends on the voter's access to information, which has been widely studied, but also the voter's capacity to process this information correctly. Based on these results, it would be reasonable to expect a lack of information to make voters worse-off. It makes them less able to scrutinise the actions of politicians, and less able to evaluate the impact of these actions on their welfare. However, additional information is not always beneficial if it affects the strategic behaviour of policy makers. Overconfidence is a special type of bias. Voters are not poorly informed per se, but think that their information is of better quality than it really is. This implies that overconfident voters are not uninformed, but lack the capacity to draw correct inferences from the information they have. This misperception was originally documented by Alpert & Raiffa (1982) who show that subjects consistently overestimates the accuracy of their predictions. Moore & Healy (2008) identify three different types of overconfidence: over-estimation, over-placement, and over-precision, and show that the last type – the one studied in this paper – is more persistent. Block & Harper (1991) suggest that overconfidence could be driven by an anchoring-and-adjustment process, while Ortoleva & Snowberg (2015) show how neglecting the correlation between different information sources leads to overconfidence. Given the prevalence of this trait, this paper addresses the question: how does the misperception of their information – rather than the lack of information – affect voters' ability to incentivise and select politicians?

I build a model in which an incumbent politicians is office motivated, can exercise costly effort that benefits the voter, and whose competence is unknown to both herself and the voter. There are two periods: in each period, the politician in office chooses how much policy effort to exert. This effort and her competence determine her performance which affects the voter's utility with some noise. Voters can therefore use their first period utility as a signal of the politician's competence. At the end of the first period, the voter decides whether to re-elect the politician based on this noisy signal. If the incumbent is not re-elected, a challenger takes office in the second period. The less noisy the information received at the end of the first period, the easier it is for the voter to draw inferences about the incumbent's competence. I am interested in the effect of the voter's overconfidence about this information. That is, the voter believes that his signal of the policy outcome at the end of the first period is less noisy (has lower variance) than it really is. Overconfidence affects equilibrium strategies through two separate channels: thinking that signals are more informative than they really are and thinking that the other player interprets the information the same way as you do. Misunderstanding the precision of the information received therefore leads the voter to mis-evaluate the equilibrium action of the politician. Politicians are typically more strategic than voters. They have more experience of the political process and dedicate more time and resources to understand the information that voters have, for example by using opinion polls. The model captures this by assuming that the politician is aware of the voter's bias. As a result, unlike other behavioural models of accountability, the politician can manipulate the voter's beliefs, even in equilibrium.

Despite the possibility of manipulating beliefs, I show that it is possible for the incumbent to exert more effort when the voter is overconfident than when she is not, and as a result to make the voter better-off. This arises because an overconfident voter will judge a politician who fails to achieve a sufficiently high performance harshly. Because the voter thinks her information is precise, she does not attribute the correct portion of a politician's unsuccessful performance to luck. The politician, aware of the voter's misperception, reacts by exerting more effort to be re-elected. Selection is always worse when the voter is overconfident because the voter mistakes a high performance for a signal of competence and is therefore more likely to re-elect incompetent types. The expected type of the second period incumbent is therefore always lower when the voter is overconfident. The first effect can sometimes dominate the second, and the voter can be overall better-off when overconfident. I characterise a condition on the level of overconfidence of the voter for the politician to exert higher effort when voters are overconfident. In particular, I show that, when the marginal benefit of effort is decreasing in effort and the difference in competence of different types of politicians is not too large, the politician exerts more effort if and only if the voter is not too overconfident.

Given that selection is always worse under overconfidence, that condition is necessary for welfare to improve. Finally, I provide parameter values that satisfy this condition and under which the overconfident voter's welfare is higher. This does not arise without overconfidence, or when the voter is aware of his overconfidence, as he would then correctly anticipates the action of the incumbent.

This therefore differs from both models with rational voters and from models with voters suffering from behavioural biases in which beliefs about the equilibrium being played do not differ.

This paper contributes to the literature on accountability by explicitly considering a natural situation in which players may not only have different information but also different second-order beliefs (their beliefs about the equilibrium strategies each player follows). It has implications for assessing the success of elections as accountability mechanisms, and for evaluating the effect of voter incompetence. First, trying to correct the bias may have unintended consequences. For instance if the bias is due to the correlation of news, which is due to concentration of ownership in the media (see Levy et al. 2018), then that concentration could actually be beneficial. Second, if voters suffer from overconfidence, increasing information (to reduce the variance of the noise) can have consequences that are hard to predict because of their effect on the voter's second-order beliefs.

Related literature

This paper extends the recent literature in behavioural political economy, by looking at the effect of a common psychological bias on a standard model of political accountability.

It contributes to the large existing literature on political accountability, started by Barro (1973), Ferejohn (1986), Austen-Smith & Banks (1989), Harrington (1993) and Persson et al. (1997). This early literature focused on the moral hazard problem that arises when politicians are able to extract socially wasteful rents from office holding and on how political institutions can mitigate these issues. It did not explicitly consider the learning process of voters who try to evaluate the quality of politicians. Fearon (1999) further developed the idea that elections can serve as both a means to align politician's incentives with voters' objectives and to select better politicians. The development and limitations of this literature is summarised in Besley (2007) and Ashworth (2012).[1]

Recent pieces of work in political economy have investigated the effect of behavioural biases on political attitudes and political behaviour. In particular, a certain stream of this literature has focused on the role of correlation neglect in politics. Ortoleva & Snowberg (2015) show that correlation neglect leads to overconfidence and that overconfidence is related to more extremism, stronger partisan identification and higher turnout. Levy & Razin (2015) show that, under certain conditions, voters with correlation neglect can aggregate information better than voters who are perfectly aware of the correlation between their information sources. Finally, Levy & Razin (2014) show that, by a similar mechanism, correlation neglect can lead to more or less platform polarisation, depending on the

[1]Other recent theoretical models of political accountability include Ashworth (2005), who develops a multi-period model of political accountability and selection where politicians choose an allocation of effort across different tasks. Meirowtiz (2007) develops an infinite horizon model of policy choice, where politicians can exploit information asymmetries to implement policies closer to their ideal points, subject to the constraint of re-election. Snyder & Ting (2008) add interest groups to models of electoral accountability. Padro i Miquel & Snowberg (2012) develop a model with both elections and primaries, and look at the politician's decision to implement or not the party's agenda. Bonfiglioli & Gancia (2013) show how political myopia can arise in a model of accountability when politicians need to choose between short-run and long-run policies. Finally, Bidner & Francois (2013) introduce norms in society that govern how to reward or punish politicians, and evaluate the behaviour of politicians who face a choice to transgress or not these norms.

competitiveness of the electoral system.

The first two of these three papers focus exclusively on voters' attitudes and behaviour. The third one introduces some strategic response from politicians to voters' biased beliefs through their choice of electoral platform. However, elections do not only offer a way to choose policies, but also to choose politicians. This paper aims to build on the literature on correlation neglect and overconfidence by applying it to models of political accountability and selection.

This paper is most closely related to a literature that evaluates the effect of behavioural biases on political accountability.[2] Some papers study the role of voters biases without taking into account strategic behaviour. For instance, Bendor et al. (2010) evaluate different retrospective voting rules that voters may follow, and examine the effect of voters' mis-perceptions of the political outcome. Kappe (2013) extends this model to asses the impact of loss-averse voters on political accountability. In both of these models however, voters follow a simple, non-strategic, adaptively rational behaviour, and politicians do not respond strategically to voters' actions. Lockwood (2017) evaluates the effect of confirmation bias on pandering. The model is based on the pandering model of Maskin & Tirole (2004), and treats confirmation bias as voters misinterpreting a binary signal. The author shows that under some conditions, confirmation bias can increase welfare by reducing the incentives to pander. However, in the version that is comparable to the present model, the author shows that confirmation bias always reduces voter's welfare. Ashworth & Bueno De Mesquita (2014) investigate whether political competence (defined as either more information or lack of behavioural bias) always leads to a higher welfare for voters, when taking into account the strategic reaction of politicians to voters' information. In particular, they show that in some cases, voter welfare can be higher when voters are behavioural, if the behavioural bias from which voters suffer leads them to change their optimal reelection threshold. The authors look in particular at the case of voters who fail to filter (i.e. who attribute responsibilities to politicians for events that are out of the politicians' control). They show that in this case, failure to filter can lead to higher welfare. Their main model is a two-period policy choice model, but the appendix includes a two-period effort choice

[2] The idea that behavioural biases may distort the incentives that political institutions are expected to provide has been mentioned in previous works. For instance, with reference to models of political accountability, Besley (2007) suggested that "Going forward it would be interesting to understand better what the differences are between behavioral models of politics and the postulates of the strict rationality supposed here. It would be useful to understand when simple and sensible behavioral rules lead to large policy distortions." (Besley 2007, p. 132). More recently, Ashworth (2012) noted that "Recent work in psychology and economics provides resources for modeling information processing in ways that deviate from the Bayesian rational expectations standard; incorporating these factors into political agency models, with and without third-party monitoring, will be an important challenge in the future."

model, which is the basis of the model used in this paper. Overconfidence differs from failure to filter in two ways. First, overconfidence involves a mistaken reduction in the variance of the noise in the incumbent's performance, while failure to filter increases this variance. Second, overconfidence is a genuine mistake, in the sense that the voter's utility will not have the distribution that the voter thinks it has. With failure to filter, by contrast, the voter will be affected by the additional shock to utility. The mistake there is to fail to adjust for *observable* noise. If this extra noise was not observable, the voter would not be making any mistake, the whole accountability problem would simply be noisier. Because the beliefs of the voter and the incumbent over the utility of the voter do not diverge, assuming that each player has correct higher order beliefs is reasonable. This generates important differences between the model presented here and that of Ashworth & Bueno De Mesquita (2014). In contrast to other types of biases that have been recently studied in this context (Ashworth & Bueno De Mesquita (2014), Lockwood (2017)), the combination of mistaken beliefs about the distribution of payoffs and of infinite action and payoff spaces can result in systematic mistakes by the voter about the equilibrium actions of the politician. The differences between the results presented here and those of Ashworth & Bueno De Mesquita (2014) and Lockwood (2017) are evaluated in more details in section 3.4.

The rest of the paper is organised as follows. Section 3.2 describes the setup of the model. Section 3.3 characterises the choice problem faced by the two players in a generic setup, defines appropriate restrictions on higher order beliefs and solves for the equilibrium beliefs and actions of the two players. Section 3.4 evaluates the differences between overconfident voters and voters with correct beliefs in a special case of the model. Section 3.5 concludes by discussing the implications of these results and comparing them with those of other recent studies. All proofs are relegated to the appendix.

3.2 Model

Environment

The model has one incumbent politician, a pool of potential challengers and a representative voter. The politician's competence can take two values $\theta \in \{\theta_L, \theta_H\}$ with $0 < \theta_L < \theta_H$. The politician's type is unknown to both the voter and the politician herself, but both share the same interior prior belief that the politician's competence is high $\mathbb{P}(\theta = \theta_H) = p,\ p \in (0, 1)$. I denote by θ_I the competence of the incumbent and θ_C

the competence of the challenger.

There are two periods. In the first period, the incumbent politician chooses a level of effort $a_1 \in \mathbb{R}$. At the end of the first period, the voter receives utility $u_1(a_1, \theta)$, update her beliefs about the type of the incumbent politician and decides whether to re-elect the incumbent. Let $r \in \{0, 1\}$ the re-election rule of the voter, where $r = 0$ means 'not re-elect' and $r = 1$ means 're-elect'. In the second period, the politician (the incumbent, if re-elected, or the challenger, drawn randomly from the pool of available politicians) chooses again an effort level $a_2 \in \mathbb{R}$, and the voter receives utility $u_2(a_2, \theta)$.

The outcome in each period is a function of the politician's type and effort and is denoted $f(a_t, \theta)$. I make the following assumptions on this function:

Assumption 3. *The outcome function satisfies: (1)* $\left|\lim_{a\to 0} f(a, \theta)\right| < +\infty$*, (2)* $\frac{\partial f(a,\theta)}{\partial a} \geq 0$ *and* $\frac{\partial f(a,\theta)}{\partial \theta} \geq 0$*, and (3)* $\lim_{a\to 0} \frac{\partial f(a,\theta)}{\partial a} > 0$ *and* $\lim_{a\to +\infty} \frac{\partial f(a,\theta)}{\partial a} < +\infty$*.*

That is, the outcome increases in effort a and is bounded when effort tends to 0. The marginal return to effort is strictly positive at low levels of effort and is bounded above as effort tends to infinity. The outcome is also increasing in the politician's competence θ.

The voter's per period utility depends on the political outcome $f(a_t, \theta)$ in that period and on some random utility shock ε_t:

$$u_t(a_t, \theta) = f(a_t, \theta) + \varepsilon_t \tag{3.1}$$

The random variable ε_t follows an 'objective' distribution with CDF G, density g, support $\mathbb{R}$, mean 0 and variance σ^2. In particular, I assume that the utility shock follows a normal distribution:

Assumption 4. *The shock* ε_t *is distributed as* $\varepsilon_t \sim N(0, \sigma^2)$*.*

The issue of overconfidence arises because the voter believes that the variance of the random shock ε_t is lower than its actual variance. Let $\sigma_V^2 < \sigma^2$ the variance believed by the voter. This form of overconfidence corresponds to the 'over-precision' described in Moore & Healy (2008) or as defined in Koehler & Harvey (2008). The distribution G is therefore objective in the sense that, should the voter not suffer from overconfidence, he would agree that this is indeed the distribution of ε_t. The voter's prior over ε_1 is therefore $\varepsilon_1 \sim N(0, \sigma_V^2)$. I denote the voter's belief about the distribution of the shock by G_V with density g_V.

The voter cannot observe the effort choice or the competence of the politician. However, his utility at the end of period 1 constitutes a noisy signal of the political outcome, from which he can draw inferences about the politician's competence.

Utilities

Given his per-period payoff and information, the voter's payoff is captured by the following two-period discounted expected utility function:

$$U_v = \mathbb{E}[u_1(a_1, \theta_I)] + \delta\big[r(u_1)\mathbb{E}[u_2(a_2, \theta_I)|u_1] + (1 - r(u_1))\mathbb{E}[u_2(a_2, \theta_C)]\big] \tag{3.2}$$

Where $\delta \in (0, 1)$ is a discount factor, and expectations are taken over the politicians' competences and the utility shock.

The politician derives utility from holding office, which is normalised to 1, and disutility from exerting effort captured by the function $c(a)$. The incumbent's utility function is therefore:

$$U_I = [1 - c(a_1)] + \delta\mathbb{E}[r(u_1)][1 - c(a_2)] \tag{3.3}$$

And the challenger's utility is:

$$U_C = \delta(1 - r(u_1))[1 - c(a_2)] \tag{3.4}$$

I make the following assumption on the cost of effort:

Assumption 5. *The cost of effort $c(a)$ is increasing and strictly convex in a, $c(0) = 0$ and $c'(0) = 0$.*

Equilibrium concept

Different choices of equilibrium concept can lead to different predictions on the action choices of the two players and the ultimate outcomes. In particular, requiring players to have correct beliefs about the equilibrium strategies of other players yields the prediction that overconfident voters correctly anticipate the effort of the incumbent politician, set their re-election threshold accordingly, and that the incumbent exerts the same level of effort independently of the degree of overconfidence of the voter (including when voters suffer from no overconfidence at all).

This approach imposes a somewhat unintuitive restriction on higher-order beliefs. Namely, it requires the voter to either believe that the politician's behaviour is optimising with respect to incorrect beliefs regarding the distribution of the random shock ε

(incorrect from the point of view of the voter, though correct compared to the objective distribution), or to believe that the politician is not rational. In addition, it requires the voter to conjecture that the politician is choosing this equilibrium strategy, despite being unable to observe either the true distribution of the random shock or the action of the politician. This means that the voter should have correct beliefs over the equilibrium action choice of the politician $a_1 \in \mathbb{R}$, despite having incorrect beliefs about the objective distribution of the utility shock. In other words, the voter would have correct second-order beliefs about the first-order beliefs of the politician[3], but incorrect (marginal) first-order beliefs (over the payoff structure).

A more intuitive approach would be to only restrict the players to have common knowledge of the rationality of other players, given their beliefs, and to have beliefs that are consistent, given an exogenously determined feedback partition (i.e. that their beliefs are consistent with what they are able to observe, given the equilibrium strategies of all players). Esponda (2013) developed an equilibrium concept, Rationalisable Conjectural Equilibria (RCE), which imposes exactly these restrictions. I use the framework developed in that paper to derive the equilibrium strategies of the players.

3.3 Equilibrium strategies

In this section, I first derive some results that are independent of the voter's higher-order beliefs and his misperceptions about fundamentals of the game. I then discuss the issues that can arise from the voter's misperceptions and characterise the equilibrium strategies given these misperceptions.

3.3.1 Second period effort and voter's problem

The politician's strategy in the second period is independent of the voter's beliefs. Given this strategy, I can then describe the problem faced by the voter at the end of the first period, and that faced by the incumbent in the first period.

From the utility functions above, and given that second-period effort is not contractible, it is easy to see that any type of politicians will exert no effort in the second period: $a_2^* = 0$. Given that $\frac{\partial f(a_2,\theta)}{\partial \theta} > 0$, the voter will re-elect the incumbent if her expected type is higher than the challenger's expected type given the information the voter obtained in the first period. In addition, since the voter cannot commit to a re-election rule ex ante, the voter cannot directly influence the politician's first period effort through the choice of re-election

[3]More precisely, over the marginal first-order beliefs of the politician over the payoff structure.

rule. These two observations imply that the voter's re-election strategy follows a threshold rule, in which the voter re-elects the politician if and only if her utility is above a certain level. Let $\bar{u}$ the threshold on the voter's first period utility above which the incumbent is re-elected. These observations are summarised in the following proposition.

Proposition 14. *Second period effort and re-election rule.*

1. *In the second period, the politician in office exerts zero effort* $a_2 = 0$.
2. *The voter correctly believes the politician will choose this action independently of his beliefs on the politician's competence.*
3. *The voter re-elects the incumbent if and only if his first-period payoff is greater than a threshold* $\bar{u}$, *and elects the challenger otherwise.*

The voter follows a threshold strategy as long as the distribution of the shock satisfies the monotone likelihood ratio property (MLRP) which the normal distribution does. All these results hold as long as there is common knowledge that the game terminates after the second period, common knowledge of rationality, and common knowledge that the politician's competence is unknown to all players. Proposition 14 is therefore satisfied for any belief hierarchies of the voter and the politician that satisfies these assumptions.

Given these results, the game can be reduced to a one-period game. The incumbent chooses a level of effort, the voter observes her payoff at the end of period 1 and re-elects the incumbent if that payoff is above the threshold. The threshold is determined by three factors: the voter's beliefs over the types of incumbent and challenger politicians, the performance function $f(a, \theta)$, and the voter's conjecture of the incumbent's equilibrium choice of effort in period 1. The first two factors are common knowledge and the voter has no mis-perceptions over those, but the third factor will be affected by the voter's mis-perceptions. However, given some arbitrary re-election threshold $\bar{u}$, the incumbent's choice of effort is independent of the voter's mis-perception. In other words, the actual first period effort is only affected by the voter's mis-perception through the threshold. This effort choice maximises the probability of being re-elected subject to the cost of effort. Given that the voter re-elects the incumbent if $u_1(a_1, \theta) > \bar{u}$, and that $u_1(a_1, \theta) = f(a_1, \theta) + \varepsilon_1$, the probability of re-electing the incumbent is:

$$\mathbb{P}(r = 1) = \mathbb{P}\left(\varepsilon_1 > \bar{u} - f(a_1, \theta)\right) = 1 - G\left(\bar{u} - f(a_1, \theta)\right) \tag{3.5}$$

The incumbent therefore solves:

$$\max_{a_1} \ \delta\Bigg[1 - pG\Big(\bar{u} - f(a_1,\theta_H)\Big) - (1-p)G\Big(\bar{u} - (f(a_1,\theta_L)\Big)\Bigg] - c(a_1)$$

Let $V(a_1, \bar{u}) = \delta\left[1 - pG\left(\bar{u} - f(a_1,\theta_H)\right) - (1-p)G\left(\bar{u} - (f(a_1,\theta_L)\right)\right]$. The first-order condition of this maximisation problem is then,

$$c'(a_1^*) = \frac{\partial V(a_1,\bar{u})}{\partial a_1}\Big|_{a_1=a_1^*} \tag{3.6}$$

The second-order condition is satisfied if the cost function is sufficiently convex. In particular, it needs to satisfy

$$c''(a_1) \geq \frac{\partial^2 V(a_1,\bar{u})}{\partial a_1^2}, \quad \forall\ a_1 \in \mathbb{R}^+ \tag{3.7}$$

3.3.2 First-period effort and re-election threshold

I now turn to characterising the re-election threshold. This threshold depends on the player's higher-order beliefs. That is, their beliefs about the other player's strategies and beliefs. To clarify how these restrictions affect equilibrium strategies, I consider two cases. In the first case, the players agree to disagree. That is, the voter believes that the shock ε_t follows a normal distribution with variance σ_V^2, but knows that the incumbent believes the variance of that shock is σ^2. In the second case, the voter is naive. He thinks that the shock has variance σ_V^2 and believes that the politician shares the same beliefs. The politician is aware of the voter's naivety: she knows the shock has variance σ^2 and knows that the voter believes that variance is σ_V^2. In each case, a different equilibrium concept is required to ensure that beliefs hierarchy are consistent. Weak perfect Bayesian equilibrium (WPBE) is appropriate for the first case. This is the equilibrium concept used in most models of accountability (e.g. Ashworth & Bueno De Mesquita 2014). In the second case, I follow the rationalisable conjectural equilibrium (RCE) proposed by Esponda (2013). I characterise the equilibrium strategies in both cases and then show how the first case (agreeing to disagree) requires some unintuitive assumptions about the players' beliefs. I derive some conditions for existence and uniqueness for both cases.

Strategies when the voter and the incumbent agree to disagree

In a weak perfect Bayesian equilibrium, players' actions maximise their expected utility given their beliefs, and beliefs are updated according to Bayes's rule. In equilibrium, players should have correct beliefs about other player's actions. Requiring players to have correct beliefs about other player's equilibrium actions leads to a straightforward conclusion in this model: overconfidence has no effect on the incumbent's choice of effort or on the selection of politicians.

To see this, I derive the re-election threshold of the voter. First recall that the voter re-elects the incumbent if, given the first period payoff she observes, she believes the incumbent is more competent than a randomly selected challenger. Let $\Phi(x)$ denote the CDF of the standard normal distribution, and $\phi(x)$ its density. Under assumption 4 $\frac{\varepsilon_1}{\sigma_V} \sim N(0,1)$. In addition, let $\tilde{a}_1$ denote the belief of the voter over the politician's equilibrium choice of effort, then the voter re-elects the incumbent if:

$$\frac{p\phi\left(\frac{u_1 - f(\tilde{a}_1,\theta_H)}{\sigma_V}\right)}{p\phi\left(\frac{u_1 - f(\tilde{a}_1,\theta_H)}{\sigma_V}\right) + (1-p)\phi\left(\frac{u_1 - f(\tilde{a}_1,\theta_L)}{\sigma_V}\right)} \geq p \tag{3.8}$$

The left-hand side of this inequality is increasing in u_1 given that ϕ satisfies the MLRP. The voter therefore re-elects the politician if and only if her first period payoff is higher than the value of u_1 for which inequality (3.8) holds with equality. Re-arranging, and using the symmetry of the normal distribution, the threshold must solve:

$$\bar{u} = \frac{f(\tilde{a}_1,\theta_H) + f(\tilde{a}_1,\theta_L)}{2} \tag{3.9}$$

The final step is to solve for the voter's conjecture about the politician's effort $\tilde{a}_1$. If the voter is aware that the incumbent has different beliefs regarding the distribution of the random shock, then the voter's conjecture will be based on the incumbent's belief about that distribution. Indeed, at the point of choosing her effort level, the politician anticipates that she will be judged based on that conjecture and the realised shock. She therefore chooses effort based on her beliefs about that shock. In turn, the voter expects the politician to follow this decision process and her conjecture is therefore based on the incumbent's beliefs: her conjecture solves condition 3.6, which is based on the objective distribution G.

Combining equations (3.6) and (3.9) gives the following characterisation of the players' strategies

Proposition 15. *The weak perfect Bayesian equilibrium strategy profile satisfies*

- *The incumbent's action a_S^P solves*

$$c'(a_S^P) = \frac{\delta}{\sigma}\phi\left(\frac{f(a_S^P,\theta_H) - f(a_S^P,\theta_L)}{2\sigma}\right) \times \left[p\frac{\partial f(a_1,\theta_H)}{\partial a_1}\Big|_{a_1=a_S^P} + (1-p)\frac{\partial f(a_1,\theta_L)}{\partial a_1}\Big|_{a_1=a_S^P}\right] \tag{3.10}$$

- *The voter's cutoff is given by* $\bar{u} = \frac{f(a_S^P,\theta_H)+f(a_S^P,\theta_L)}{2}$

This equilibrium exists given assumptions 3 and 5 and it is unique if the incumbent's objective function $V(a_1,\bar{u})$*, evaluated at the equilibrium threshold* $\bar{u}$ *is strictly concave in* a_1*.*

Note that this equilibrium only depends on the incumbent's beliefs about the distribution of the random shock, not on those of the voters: σ_V does not appear in equation (3.10). Therefore, the resulting level of effort, quality of selection and voter welfare is the same whether the voter is overconfident or not. Intuitively, the incumbent chooses an action that is optimal given her (correct) prior beliefs about the random utility shock ε. Given this behaviour, the voter's best response is to choose the re-election threshold defined by (3.9).

This result requires some demanding assumptions on the higher-order beliefs of the players. Despite the fact that the voter cannot observe the actions of the politician, and that the feedback that the voter receives does not allow her to verify that her conjecture about the politician's action is correct, the voter's actions (the re-election threshold) is based on the correct equilibrium play of the politician. In addition, the voter has to believe that either the politician is not rational or to agree to disagree with her: believe that the politician has a different prior, but continue to believe her prior is correct. Assuming that players agree to disagree makes some demanding assumptions given the focus on overconfidence. It requires the voter's second-order belief over the incumbent's prior beliefs to be correct, while his own prior beliefs are incorrect.

Strategies when the voter is naive

Because this model explores the effect of voters' systematic mistakes, a more intuitive approach is to assume that the voter is unaware that the politician has a different prior belief. This requires some different restrictions on higher-order beliefs in order to maintain the assumption of common belief of rationality. In particular, the requirement that players have correct beliefs over the equilibrium strategy of the other players needs to be relaxed. To explore this question, I use the framework of rationalisable conjectural equilibria developed by Esponda (2013).

In particular, I look instead for equilibria in which players do not have correct beliefs over the equilibrium strategies of the other players.[4] Following Esponda (2013), a strategy profile is a rationalisable conjectural equilibrium if there exists a belief space $\mathcal{B}$ such that players maximise their expected utility, given their beliefs, and these beliefs are consistent given a partition, and such that there is common knowledge of rationality and consistency.

To find the set of such equilibria, I first define the exogenous feedback partition that each player is facing, based on the timing and information structure of the model described in section 3.2. I then find the set of strategy profiles that satisfy rationality and consistency, and finally, I refine this set of equilibria by imposing restrictions on higher-order beliefs, in particular, common knowledge of rationality and consistency.

Given that the voter only observes her utility at the end of the first period, and at the end of the second period, I restrict the support of the beliefs over the epistemic state that generates the incumbent's actions and the distribution of the utility shock ε to be consistent with the observed utility. Since both the incumbent's action and the random shock have support over the entire real line, this restriction is not very stringent: any observed utility is consistent with some combination of incumbent effort and some realisation of the random shock. In addition, I make the natural restriction that the voter should know her own action (or her choice of re-election threshold). I also maintain the assumption that the voter knows the correct distribution of the politician's type. I restrict the incumbent to know with certainty the 'correct' (objective) distribution of both the random shock, and of her own type. The incumbent also observes her own actions, and observes whether or not she is re-elected. This assumption does not place strong restrictions on the beliefs of the incumbent, since re-election could be consistent with an infinity of combinations of re-election thresholds and realisation of the random shock.

I restrict the voter to believe that the incumbent has the same beliefs as the voter over the equilibrium strategies, and the fundamentals of the game. This captures the idea that the voter is naive: she is not aware that she is mistaken over the distribution of the random shock, and therefore believes that the incumbent is facing the same fundamental game as she is.

[4]But in which players have common belief of rationality (they believe that other players are maximising their own payoffs, given the beliefs that they think other players have) and have a hierarchy of beliefs consistent with the feedback they can obtain from their payoffs or other signals about other players' strategy.

Given these beliefs, the voter believes that the incumbent solves the following problem:

$$\max_{a_1} \ \delta\left[1 - pG_V\Big(\bar{u} - f(a_1, \theta_H)\Big) - (1-p)G_V\Big(\bar{u} - (f(a_1, \theta_L)\Big)\right] - c(a_1) \tag{3.11}$$

For a given threshold $\bar{u}$.

Let $V_V(a_1, \bar{u}) = \delta\left[1 - pG_V\left(\bar{u} - f(a_1, \theta_H)\right) - (1-p)G_V\left(\bar{u} - (f(a_1, \theta_L)\right)\right]$. If the voter believes that the random shock ε is distributed as $\varepsilon \sim N(0, \sigma_V^2)$, then the voter believes that the incumbent's equilibrium strategy a^V solves

$$c'(a^V) = \frac{\partial V_V(a_1, \bar{u})}{\partial a_1}\Big|_{a_1 = a^V} \tag{3.12}$$

This equilibrium choice of effort a^V, determines the re-election threshold

$$\bar{u}^V = \frac{f(a^V, \theta_H) + f(a^V, \theta_L)}{2}.$$

Finally, as the voter believes that the distribution of ε is G_V, and the voter believes that the incumbent will have correct beliefs about the equilibrium threshold, the voter believes that the incumbent's equilibrium strategy solves

$$\begin{aligned} c'(a^V) = \delta\frac{1}{\sigma_V}\phi\left(\frac{f(a^V, \theta_H) - f(a^V, \theta_L)}{2\sigma_V}\right) \\ \times\left[p\frac{\partial f(a_1, \theta_H)}{\partial a_1}\Big|_{a_1 = a^V} + (1-p)\frac{\partial f(a_1, \theta_L)}{\partial a_1}\Big|_{a_1 = a^V}\right] \end{aligned} \tag{3.13}$$

This equation has the same form as equation (3.10), except for the different standard deviation σ_V used. Therefore, the conditions for the existence and uniqueness of a^V are the same as those for a_S^P, replacing σ by σ_V.

Therefore, in an RCE, the equilibrium choice of threshold of the voter is $\bar{u}^V = \frac{f(a^V, \theta_H) + f(a^V, \theta_L)}{2}$, and the equilibrium beliefs of the voter over the incumbent's action are a^V that solves equation (3.13).

Recall that the incumbent is aware of the voter's mistakes, and is aware that the voter is unaware that the incumbent has different beliefs. Therefore, the incumbent correctly believes that the voter uses a threshold $\bar{u}^V = \frac{f(a^V, \theta_H) + f(a^V, \theta_L)}{2}$, where a^V is defined by equation (3.13), as above.

The equilibrium strategy of the incumbent is then determined by her maximisation problem given a re-election threshold based on the voter's incorrect conjectured equilibrium. The incumbent's first period $a_1 = a^I$ solves

$$c'(a^I) = p\delta\frac{1}{\sigma}\phi\left(\frac{\frac{f(a^V,\theta_H)+f(a^V,\theta_L)}{2} - f(a^I,\theta_H)}{\sigma}\right)\frac{\partial f(a_1,\theta_H)}{\partial a_1}\Big|_{a_1=a^I}$$
$$+ (1-p)\delta\frac{1}{\sigma}\phi\left(\frac{\frac{f(a^V,\theta_H)+f(a^V,\theta_L)}{2} - f(a^I,\theta_L)}{\sigma}\right)\frac{\partial f(a_1,\theta_L)}{\partial a_1}\Big|_{a_1=a^I} \tag{3.14}$$

I summarise these results in the following proposition:

Proposition 16. *If the voter is overconfident and naive, and if the incumbent is aware of the voter's overconfidence, then the following strategy profile and beliefs constitute a rationalisable conjectural equilibrium*

1. *The incumbent chooses an effort level a^I such that a^I solves equation (3.14).*
2. *The voter chooses a re-election threshold $\bar{u}^V = \frac{f(a^V,\theta_H)+f(a^V,\theta_L)}{2}$, such that a^V solves equation (3.13).*
3. *The incumbent believes that the voter uses a threshold $\bar{u}^V$, and the voter believes that the incumbent chooses a^V.*
4. *All higher-order beliefs of both players are correct.*

The existence of this equilibrium is guaranteed by the same conditions as when the voter is not overconfident. Indeed, the voter's problem is the same in both cases, simply with a different variance. As a result, a^V exists if a^P_S exists. Given some value a^P_S, the strict convexity of $c(\cdot)$ and the properties of $f(a,\theta)$ also guarantee that there exists an a^I that solves equation (3.14). The uniqueness of the equilibrium requires a more demanding condition however. While a^V is unique if a^P_S is unique, the uniqueness of a^I is not guaranteed by the strict concavity of the incumbent's objective function when the re-election threshold is based on the correct incumbent action. Instead, a sufficient condition is that the following function is concave in a given $\bar{u}^V = \frac{f(a^V,\theta_H)+f(a^V,\theta_L)}{2}$:

$$V^{RCE} = \frac{\delta}{\sigma}\left[p\Phi\left(\frac{\bar{u}^V - f(a,\theta_H)}{\sigma}\right) + (1-p)\Phi\left(\frac{\bar{u}^V - f(a,\theta_L)}{\sigma}\right)\right]$$

Proposition 17. *The equilibrium defined in proposition 16 always exists and is unique if $V^{RCE}(a)$ is concave in a for any $a > 0$ and given $\bar{u}^V = \frac{f(a^V,\theta_H)+f(a^V,\theta_L)}{2}$.*

This condition is more demanding than necessary and the equilibrium could be unique even if the condition is not satisfied. However, it is easy to verify for a given set of parameters, and I show that it is satisfied for a given functional form and given parameters in the next section.

Assuming that the voter is naive therefore implies that the voter's perception of the incumbent's first-period effort and the actual first-period effort differ in equilibrium. As long as $\sigma^V \neq \sigma$, then $a^V \neq a^I$, and there is a possibility for the politician to take advantage of the voter's mis-perception. This creates a wedge between the equilibrium effort perceived by an overconfident voter and the one perceived by a fully rational voter. In turn, this leads to a different re-election rule, and therefore a different expected second-period politician type in equilibrium, than if the voter suffered no overconfidence. Given these equilibrium beliefs and strategies, the effect of overconfidence on voter welfare arises through the difference between the voter's expectation of the politician's optimal action and the politician's actual optimal action (since the politician does not suffer from overconfidence, and has correct beliefs over the mistakes of the voter).

I now turn to the consequences of the voter's overconfidence and naivety on the incumbent's effort, the voter's capacity to select high quality politicians and the voter's welfare in a special case of the model.

3.4 Comparing the welfare of overconfident voters and of voters with correct beliefs

In this section, I compare the equilibrium politician effort, politician selection and voter welfare when the voter is overconfident and naive and when the voter is fully rational. Because of the voter's naivety, the politician is able to manipulate the voter's beliefs in the sense that she can choose a level of effort different than the one the voter conjectures, and as a result induce different beliefs about her type. Despite this possibility, I show in a simple case that the voter can be better off when overconfident.

To keep the comparisons tractable, I assume that the outcome function $f(a, \theta)$ is linear and additively separable in effort and competence.

Assumption 6. *The performance function is linear in effort and ability $f(a, \theta) = a + \theta$.*

In addition, I normalise $\theta_L = 0$ and $\sigma^2 = 1$.

Comparing effort levels

I begin by comparing the equilibrium level of effort when the voter is overconfident and naive to the level of effort when the voter is fully rational. As shown in Proposition 15, it is the naivety of the voter that allows the difference in effort level between the case of an overconfident voter and that of a rational one. Therefore, the effort level exerted when a voter is fully rational is the same as that exerted when the voter is overconfident but aware of her disagreement with the politician. The equilibrium level of effort in the case of a rational voter, denoted a^R, is therefore the solution to the first-order condition of the politician's problem when the players agree to disagree, given in equation (3.10). That is, $a^R = a^P_S$, and given assumption 6 and the normalisation, a^R solves:

$$c'(a^R) = \delta\phi\left(\frac{\theta_H}{2}\right) \tag{3.15}$$

When the voter is overconfident, the voter's belief about the politician's effort, a^V, solves equation (3.12) and the actual equilibrium strategy of the politician, a^I solves equation (3.14). That is, a^V and a^I solve

$$c'(a^V) = \frac{\delta}{\sigma^V}\phi\left(\frac{\theta_H}{2\sigma_V}\right) \tag{3.16}$$

$$c'(a^I) = \delta\left[p\phi\left(\frac{2(a^V - a^I) - \theta_H}{2}\right) + (1-p)\phi\left(\frac{2(a^V - a^I) + \theta_H}{2}\right)\right] \tag{3.17}$$

Given Proposition 17, a^I is unique if $\delta\left[p\phi\left(\frac{2(a^V-a)-\theta_H}{2}\right) + (1-p)\phi\left(\frac{2(a^V-a)+\theta_H}{2}\right)\right]$ is decreasing in a for any $a > 0$. This function is increasing up to a threshold $\bar{a}$ and decreasing above $\bar{a}$. Therefore, a sufficient condition for the uniqueness of a^I is that $\bar{a} \leq 0$.

Lemma 4. *Under assumption 6,*

1. *a^R exists and is unique.*

2. *a^V exists and is unique.*

3. *a^I exists, and is unique provided that $\bar{a} \leq 0$, where $\bar{a}$ solves*

$$(1-p)\left(\frac{2a^V + \theta_H}{2} - \bar{a}\right)\exp\left(2\theta_H(a^V - \bar{a})\right) + p\left(\frac{2a^V - \theta_H}{2} - \bar{a}\right) = 0$$

Comparing a^R to a^I shows that the incumbent exerts more effort when facing overconfident voters than perfectly rational voters, if she is sufficiently overconfident.

Proposition 18. *Suppose assumption 6 is satisfied and $\bar{a} \le 0$, then there exists a threshold $s^*(\theta_H)$ such that the politician's effort when the voter is overconfident, a^I, is greater than her effort when the voter is rational, a^R, if and only if $\sigma_V > s^*(\theta_H)$ (provided that σ_V is such that $\bar{a} \le 0$).*[5] *In addition, $s^*(\theta_H) \in [0,1]$ if $\theta_H < 2$ and $s^*(\theta_H)$ is increasing in θ_H.*

To understand the relationship between the level of overconfidence and the ranking of effort levels, first note that the politician's equilibrium level of effort when the voter is overconfident is always between the beliefs of the overconfident voter about the politician's effort a^V and the equilibrium effort when the voter is rational a^R. That is, either $a^R < a^I < a^V$ or $a^V < a^I < a^R$.

To illustrate the intuition behind this result, suppose for instance that $a^R < a^V$. The first observation is that, since the marginal cost of effort is increasing in effort, the marginal cost of effort when the voter is rational $c'(a^R)$ must be lower than the marginal cost of effort perceived by an overconfident voter $c'(a^V)$.

In addition, the marginal cost of effort when the voter is rational, $c'(a^R)$ must be equal to the marginal benefit of effort of a politician facing an overconfident voter, if the politician were to choose the level of effort expected by that overconfident voter. That is, if $a = a^V$, then the marginal benefit, $B(a) = \delta \left[p\phi\left(\frac{2(a^V-a)-\theta_H}{2}\right) + (1-p)\phi\left(\frac{2(a^V-a)+\theta_H}{2}\right)\right]$ satisfies:

$$\begin{aligned} B(a^V) &= \delta \left[p\phi\left(\frac{2(a^V-a^V)-\theta_H}{2}\right) + (1-p)\phi\left(\frac{2(a^V-a^V)+\theta_H}{2}\right)\right] \\ &= \phi\left(\frac{\theta_H}{2}\right) = c'(a^R) \end{aligned} \tag{3.18}$$

This is because both the politician and a rational voter use the correct variance $\sigma^2 = 1$ when computing this marginal benefit, and $a^I = a^V$ means that the marginal benefit is evaluated as if the voter correctly anticipates the politician's effort level, which is what a rational voter would do.

Combining these two observations implies that the marginal benefit of effort if the politician were to take the effort level expected by an overconfident voter $B(a^V)$ must be lower than the marginal cost at that effort level $c'(a^V)$, since $B(a^V) = c'(a^R) < c'(a^V)$. The politician would therefore prefer to reduce her effort level, and the equilibrium effort level a^I must be lower than a^V.

Moreover, when $\bar{a} \le 0$, the marginal benefit of effort of a politician facing an overconfident voter $B(a)$ is decreasing in a. This implies that the marginal benefit evaluated at the

[5] I show in Proposition 20 that the set of parameters such that both $\sigma_V > s^*(\theta_H)$ and $\bar{a} \le 0$ hold simultaneously is non-empty.

rational voter's expected level of effort a^R must be higher than at the overconfident voter's expected level: $B(a^V) < B(a^R)$. As a result, the marginal benefit of effort of a politician facing an overconfident voter, evaluated at the level of effort expected by a rational voter $B(a^R)$ must be greater than the marginal cost at that level: $B(a^R) > B(a^V) = c'(a^R)$. So the politician would like to increase her effort at this level and we must have a^I greater than a^R.

Taken together, these imply that if $a^R < a^V$, then $a^R < a^I < a^V$. A similar intuition implies that if $a^V < a^R$, then $a^V < a^I < a^R$. As a result, the ranking of effort between a situation where the voter is rational, and one in which the voter is overconfident is fully determined by the ranking of the effort levels expected by a rational voter, a^R, versus an overconfident voter, a^V.

Given this relationship, I now describe how the ordering of the effort levels expected by a rational voter, a^R, versus an overconfident voter, a^V, depends on the level of overconfidence.

Note that the marginal benefit of exerting more effort for the incumbent, from the point of view of an overconfident voter is $\frac{\delta}{\sigma_V}\phi\left(\frac{\theta_H}{2\sigma_V}\right)$. This marginal benefit is non-monotonic in the perceived variance σ_V^2. If that variance is sufficiently low, then the marginal benefit of effort is increasing in variance, since the more noise there is, the more the politician can jam the signal using her effort. Intuitively, as the variance goes to zero, the voter would be able to perfectly infer the type of the politician, if she had the correct effort conjecture.

The voter believes that she is making the correct conjecture, and therefore believes that effort is lower when the variance is σ_V^2 than when it is $\sigma^2 = 1$. As the perceived variance increases, the perceived marginal benefit becomes higher and eventually, the overconfident voter believes that the politician exerts more effort than the rational voter does: $a^V > a^R$. When perceived variance is high, the marginal benefit of effort starts decreasing in effort and the equilibrium effort perceived by an overconfident voter, a^V, eventually tends to that perceived by the rational voter, a^R. This is illustrated in figure 3.1.

Therefore, if the perceived variance is sufficiently high, $\sigma_V > s^*(\theta_H)$, an overconfident voter overestimates the marginal benefit of exerting more effort, and therefore expects the incumbent to exert more effort than a fully rational voter would, so $a^V > a^R$. As a consequence, if the incumbent were to exert the same effort as she would when the voter is rational, $a = a^R$, the marginal benefit of effort would outweigh the marginal cost, and the incumbent would be better-off increasing her effort. So we must have $a^I > a^R$ in equilibrium.

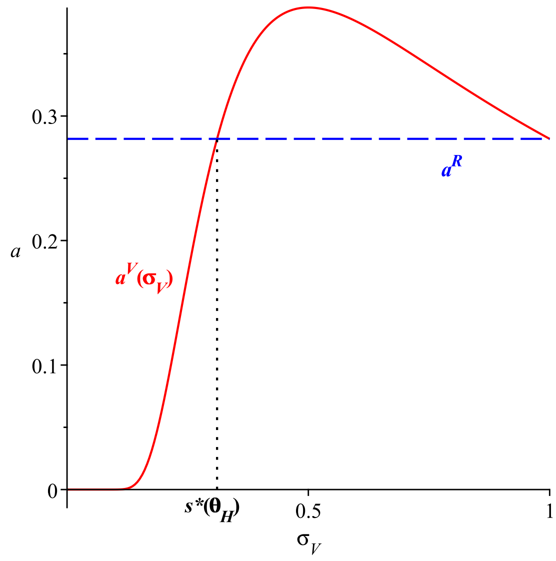


Figure 3.1: Equilibrium effort when voter is rational a^R and perceived effort when voter is overconfident a^V

When the difference in competence between different types of politicians increases (that is θ_H increases), that threshold increases too. Intuitively, a higher level of competence of the good type θ_H decreases the marginal benefit of effort perceived by both an overconfident and a rational voter. This decrease is more important for the overconfident than the rational voter, and the ordering of a^R and a^V can be reversed. To compensate, the overconfident voter needs to be less overconfident to guarantee that effort is higher than with a rational voter, so $s^*(\theta_H)$ needs to increase.

Finally, note that, conditional on $\bar{a} \leq 0$, neither the prior beliefs on the politician's type p, nor the discount factor δ affects this ordering. When voters and politicians have the same beliefs about the noise, the incumbent fully internalises the effect of a higher probability of being a high-type, p on the voter's decision. Therefore, in equilibrium, this parameter does not affect the incumbent's choice of effort. This implies that neither a^R nor a^V are affected by p. Since only the ordering of these two variables matters to determine whether effort is higher under overconfidence, p does not affect this result. Similarly, a change in δ affects both a^R and a^V in the same way, so it does not change their ordering.

These parameter matters can affect whether $\bar{a} \leq 0$, however. As the overconfident voter does not fully internalise the effect of changes in these parameters, the marginal benefit $B(a)$ of the politician facing an overconfident voter depends on the prior p and the discount rate δ. Recall that $\bar{a}$ is the point above which the marginal benefit of effort of an incumbent facing an overconfident voter is decreasing. A higher p means the voter is more likely to face a competent politician, who is more likely to produce a high outcome, all else equal, and as a result benefits less from exerting higher effort. Therefore, $\bar{a}$ is more likely to be below 0 if p is high. By contrast, a higher δ raises the marginal benefit perceived by another confident, and therefore the effort that voter expects the politician to exert a^V. In turn, this raises the threshold of re-election, and increases the marginal benefit of higher effort. Therefore, $\bar{a}$ is more likely to be below 0 if δ is low.

Politician selection

The selection of high type politicians is another concern of accountability. I evaluate the capacity of the voter to select the right politician through the ex-ante expected utility of the voter in the second period. Since the equilibrium action in the second period is $a_2 = 0$ for both types of politicians, this is equivalent to the expected type of the second period politician. For a given threshold $\bar{u}$, this is given by

$$\begin{aligned} S(\bar{u}) &= \mathbb{E}_{\theta_I}\left[\mathbb{P}(u > \bar{u} \mid \theta_I)\theta_I + \mathbb{P}(u < \bar{u} \mid \theta_I)\,\mathbb{E}[\theta_C]\right] \\ &= p(1-p)\theta_H\left[\mathbb{P}(u > \bar{u} \mid \theta_H) - \mathbb{P}(u > \bar{u} \mid \theta_L)\right] + p\theta_H \end{aligned}$$

As a result, selection improves when the voter is overconfident if and only if

$$\mathbb{P}(u > \bar{u}^V \mid \theta_H) - \mathbb{P}(u > \bar{u}^V \mid \theta_L) > \mathbb{P}(u > \bar{u}^R \mid \theta_H) - \mathbb{P}(u > \bar{u}^R \mid \theta_L)$$

Where $\bar{u}^V = a^V + \frac{\theta_H + \theta_L}{2}$ and $\bar{u}^R = a^R + \frac{\theta_H + \theta_L}{2}$. Given assumptions 5 and 6, this reduces to

$$\Phi\left(\frac{2(a^V - a^I) + \theta_H}{2}\right) - \Phi\left(\frac{2(a^V - a^I) - \theta_H}{2}\right) > \Phi\left(\frac{\theta_H}{2}\right) - \Phi\left(\frac{-\theta_H}{2}\right)$$

The next proposition establishes that this can never occur. An overconfident and naive voter is always worse at selecting good politicians than a rational voter.

Proposition 19. *If assumption 6 is satisfied, and equilibria with and without overconfidence exist and are unique, then selection is always worse when the voter suffers from overconfidence.*

Because the voter misperceives the equilibrium effort of the politician, the overconfidence of the voter introduces a wedge $a^V - a^I$ between the threshold that the voter uses and the threshold he should be using if he was rational. As the politician is aware of this wedge, she can take advantage of it to increase her chances of re-election. This may lead to higher or lower effort than the voter expects, as shown in Proposition 18. However, in both cases, it reduces the informativeness of the payoff u_1 as a signal of the politician's competence. Intuitively, the voter misinterprets this wedge as some information about the politician's competence, when it is in fact independent of that competence.

When $a^V - a^I > 0$, the payoff shock ε has to be higher for both types to be re-elected, which reduces the probability of re-election of both types, but more so for the competent type θ_H than the incompetent type θ_L. When $a^V - a^I > 0$, re-election is possible for a lower ε for both types, but this increase in re-election probability is higher for the incompetent type θ_L than the competent type θ_H. Therefore, in both cases the relative probability of electing an incompetent type θ_L increases.

Voter's welfare

Given Proposition 19, a necessary condition for the voter's welfare to be higher under overconfidence is that the politician's effort is higher. Given Proposition 18, this requires $\sigma_V > s^*(\theta)$ if $\bar{a} < 0$. This is not sufficient, however, as the negative effect on selection could outweigh the positive effect of effort. The following proposition therefore characterises a necessary, but not sufficient, condition for the voter to be better-off with overconfidence when $\bar{a} \leq 0$.

Proposition 20. *If $\bar{a} \leq 0$ an overconfident voter can be better-off than a rational voter only if $\sigma_V > s^*(\theta)$ and $\theta_H < 2$.*

The necessary and sufficient condition cannot be expressed in a closed form and describing it therefore does not unambiguously reveal when it is or is not satisfied. It is provided in full, as a system of equations and inequalities, in the appendix. A numerical example shows that there are reasonable sets of values for which this condition is satisfied, so that voter welfare can indeed sometimes be higher under overconfidence.

This happens when the voter is not *too* overconfident ($\sigma_V > s^*(\theta)$), and therefore overestimates the politician's marginal benefit of effort relative to a rational voter. In that case, the politician prefers to exert more effort to increase her chances of re-election ($a^I > a^R$). For the higher effort to compensate for the worse selection, the effect of overconfidence on selection should be relatively small. Following the intuition of Proposition 19, this happens when the overconfident voter's perception of effort a^V is not too far from the actual effort of a politician facing an overconfident voter a^I, that is $a^V - a^I$ is relatively small.[6] That distance depends subtly on the level of overconfidence σ_V, due to the shape of the normal distribution's density function.

[6] Recall from Proposition 18 that if $a^I > a^R$, then we must have $a^V > a^I$, given $\bar{a} \leq 0$.

These results imply that the voter can be better-off when he is sufficiently overconfident that the politician wants to take advantage of that overconfidence, so that a^I is different from a^R, but not too overconfident that overconfidence causes a significant loss of information, that is, a^V cannot be too far from a^I.

3.5 Discussion

Implications

The results from the previous sections reveal that voters can be better-off when they overestimate the precision of their information, and are unaware that politicians have different beliefs. This has normative implications for the effect of educating voters on voter welfare, and empirical implications for measuring the effect of information on accountability.

Since overconfidence can result from voters misunderstanding the correlation between their information sources (Ortoleva & Snowberg 2015), making voters aware of these correlations, or reducing the source of these correlations could reduce voters' overconfidence (Levy et al. 2018). This model suggests that such actions could be misguided and result in lower voter welfare, if voters are not too overconfidence. More generally, policy attempting to better inform voters should be decided based on the size of the bias that voters suffer from, and on the voter's perceptions of what politicians know.

A number of empirical studies look at the effect of improving voter's information on accountability (see Ashworth (2012) for a survey and Bhandari et al. (2018) for a list of more recent studies). This can take the form of information on the incumbent's policy choices, which this model does not address or information about the economic environment or the impact of policies. In this model, an increase in the second type of information available to both politician and voters can be interpreted as a decrease in the actual variance (σ) of the utility shock. An increase in information provided to voters only can be interpreted as a decrease in the voter's perceived variance (σ_V), and therefore corresponds to an increase in overconfidence. This paper shows that the increase in either of these parameters can have ambiguous effects on the politician's action and on the voter's welfare. They also crucially depend on the level of sophistication of the voter, such as whether he thinks the politician shares the same beliefs. The existing literature on behavioural biases and accountability (Ashworth & Bueno De Mesquita 2014, Ashworth et al. 2018) suggests that evaluating whether democratic institutions are effective by assessing whether voters are well informed is insufficient as voters face strategic politicians. They suggest that more parameters of the political environment need to be accounted for to properly assess the relationship between these two variable. This paper complements this view by suggesting that another relevant parameter is not just what voters know but how they perceive what politicians know.

Comparison with existing behavioural models of political accountability

Ashworth & Bueno De Mesquita (2014) also find that under certain conditions, irrational voters (voters who fail to filter) can incentivise the incumbent to exert higher first-period effort. In particular, they show that, in a pandering model, a low probability of having an extreme challenger (equivalent to a high p in this model), gives the incumbent more incentives to moderate (here, exerts higher effort). As expected, this is the opposite of what happens in this model, since the bias they consider (failure to filter irrelevant noise) implies that voters have more noise, while overconfidence implies that voters think that they have less noise in their signals.

The key difference between Ashworth & Bueno De Mesquita (2014)'s model and this model, however, lies in the fact that their voters correctly predicts the equilibrium action of the incumbent. This is because the voter's payoff is genuinely affected by the shock that is unobservable from the point of view of the voters (or that they choose to ignore), but observable for the politician. In this sense, the incumbent has to adopt the erroneous beliefs of the voter when choosing their effort. Overconfidence, by contrast, is a genuine mistake, in the sense that the shock to the voter's utility has a different distribution that the voter thinks it has.

Because the beliefs of the voter and the incumbent over the utility of the voter do not diverge in Ashworth & Bueno De Mesquita (2014), each player also has correct higher order beliefs over the other player's strategy. Therefore, no difference arises between the voter's beliefs about the politician's equilibrium strategy, and the actual politician's strategy. One direct implication of this less demanding form of irrationality is that selection can either improve or worsen with a voter who fails to filter, while it always decreases with an overconfident voter. A second implication is that the relationship between irrationality and welfare is more subtle when voters are overconfident, as the changes in welfare depend not only on the direct effect of overconfidence on first-period effort, but also on the indirect effect on the voter's belief about this effort.

Lockwood (2017) looks at a model of pandering: the incumbent knows the state and chooses his action in order to improve the voter's perception of his type. The model therefore differs from mine because of the possibility for the politician to signal his type through his actions. By contrast, information about the politician's competence is symmetric in my model. My model can therefore be interpreted as a model of pandering with unobservable actions but observable payoffs, in which the politician is always dissonant: she dislikes effort but increasing effort increases her chances of re-election.

Secondly, the bias studied is different. Lockwood (2017) looks at confirmation bias, the possibility of mis-interpreting a bad signal as a good one. In that sense, confirmation bias is similar to overconfidence: a voter with confirmation bias thinks that he got the correct signal when it is actually incorrect, while an overconfident voter thinks he got a precise signal when it is actually less precise. In both cases, the politician knows the correct distribution of the signal that the voter is receiving, but the voter does not know that this is the case. The main difference is therefore that overconfidence is not partisan, in the sense that the misinterpretation of information is independent of the voter's prior. A second important difference is that, with overconfidence, the voter's misinterpretation is relevant for the inference she draws: the voter and the politician agree on what is observed but disagree about how it was generated. This implies that overconfident voters have the wrong beliefs about the politician's strategy, even in equilibrium.

In the case of unobservable actions but observable payoffs, Lockwood (2017) finds that the probability that the incumbent chooses the right action (in his case, that the dissonant politician imitates the consonant one), is always decreasing in the strength of confirmation bias. This helps selection, by making it easier for the voter to distinguish between a good and a bad politician but this improved selection is not sufficient to overcome the worst first-period action. As a result, a behavioural voter always fares worse than a fully rational voter. I find the opposite: if overconfidence is not too extreme, the politician is more likely to choose the right action, but selection is worse, and the overconfident voter is overall better off.

The difference arises because in his model, more confirmation bias unequivocally decreases the marginal benefit of pandering (i.e. exerting more effort) for the dissonant type, as it raises the chances of being re-elected when the incumbent does not pander to voters. With overconfident voters instead, the marginal benefit of increasing effort can increase because the strategy expected by the voter can differ from that actually played by the politician, so additional effort creates a more positive surprise for the voter who expected a relatively low effort level in equilibrium. Therefore, the opposite results arises because of the equilibrium difference in the perception of the politician's strategy. This highlights the importance of looking at the role of bias on second-order beliefs as well as first-order beliefs.

Appendices

A Proofs of results and extensions in Chapter 1

A.1 Commitment to confidentiality or transparency

In this section, I show that the results are not driven by the policy maker's ability to commit ex-ante to the confidentiality of her information. A standard unraveling intuition would suggest that if the information is verifiable, no disclosure of the policy maker's information would make the lobbyist believe that the policy maker observed a 'bad' realisation (e.g. Grossman (1981)). However, this does not occur here. The reason is that no type has incentives to deviate from keeping the realisation r confidential. A policy-maker who observed r_1 strictly gains from keeping her information confidential, as she obtains some informational rent from inducing π_G. A policy-maker who observed r_0 is indifferent between keeping this information confidential and making it public as π_G is the information the lobbyist would choose if she knew the policy maker had observed r_0.

Formally, I assume that the policy maker cannot lie about the information she obtained but chooses whether to disclose it. Let $\hat{r} \in \{r, \varnothing\}$ the information she chooses to disclose and consider the following modification to the timing of the game: after $r \in \{r_0, r_1\}$ is realised and observed by the policy maker only, the policy maker chooses whether to disclose the realisation $\hat{r} = r$ or to keep it confidential $\hat{r} = \varnothing$. The lobbyist then chooses π, conditional on the report $\hat{r}$.

The following result captures the intuition above. Recall from Proposition 4 that if B is sufficiently high ($B > \overline{B}$), the policy maker weakly prefers the results from her investigation to be public.

Proposition 21. *In the modified game, if $B < \overline{B}$, there is an equilibrium in which:*

1. *The equilibrium choice of p is as defined by Proposition 3*
2. $\pi = \pi_G$
3. *The policy maker does not disclose her information*

If $B > \overline{B}$, it is an equilibrium for the policy maker to choose the most informative p, disclose her information, and for the lobbyist to choose his optimal public persuasion strategy as defined by Lemma 1.

Therefore, the results established in previous sections remain equilibria under this alternative assumption.[7] Allowing the policy maker to choose the confidentiality of her signal after it is realised also gives rise to other equilibria. In the case where $B < \overline{B}$, it is also an equilibrium for both types to disclose their information, or for type r_0 to disclose and type r_1 not to. Both of these equilibria can be supported by the lobbyist having off-equilibrium beliefs putting probability 1 on the deviating type being r_1.[8] When $B > \overline{B}$, there is an equilibrium in which type r_0 discloses her type and type r_1 does not. This is equivalent to full disclosure in the sense that the act of not disclosing fully reveals the type.

This points to an important difference between the mechanism described in this paper and alternative explanations for confidentiality in government. It could be suggested that information is often not disclosed because it makes a policy maker look bad. This is not explicitly modeled here, but could be an alternative way to rationalise the examples mentioned in the introduction. If this were the case, however, the policy maker would be indifferent between disclosing bad information or not, as no disclosure would amount to revealing her bad signal (bad being interpreted as r_1, or information that suggests the lobbyist is right). In this case, no disclosure would be observed in conjunction with (1) the government fully using their expertise, and (2) the lobbyist providing the minimum level of evidence. The present model suggests by contrast that the policy maker would be indifferent between disclosing *good* information or not, and would strictly prefer not to

[7] Note that these equilibria survive standard refinements such as the intuition criterion (Cho & Kreps (1987)). In the case where $B < \overline{B}$, this is straightforward as deviating from the equilibrium strategy of no disclosure requires full revelation of the signal. Therefore, the off-equilibrium beliefs of the lobbyist following a deviation are uniquely defined. In the case where $B > \overline{B}$, the reason is that type r_0 would never strictly gain by deviating for any beliefs of the lobbyist. Therefore observing no disclosure would induce the lobbyist to believe that the policy maker observed r_1 with probability 1. As a result, the policy maker would not gain from deviating to no disclosure following r_1.

[8] This equilibrium is dominated for the policy maker (by Proposition 3). Therefore selecting the policy maker-preferred equilibrium would leave only the equilibrium where information is not disclosed.

disclose bad information. No disclosure of information would be observed in conjunction with high amount of evidence presented by lobbyists.

A.2 Availability of more complex preliminary investigations

In the analysis, I restrict the policy maker to a preliminary investigation producing at two most two signals. Given that the policy maker's expertise is limited, it is reasonable to restrict attention to simple investigation that can produce only one of two conclusions, such as a memo summarising the key pros and cons of a policy and offering one policy recommendation.

In addition, when information is transparent, the policy maker cannot gain from a preliminary investigation producing more than two possible realisations. Any other realisation would induce a belief μ_{r_*} in the interval $[\underline{\mu}, \bar{\mu}]$. If $\mu_{r_*} < \frac{1}{2}$, then the lobbyist would respond with a persuasion strategy π_{r_*} and on average, the policy maker would not gain from that additional signal. $\frac{1}{2} < \mu_{r_*} < \bar{\mu}$, then the lobbyist would not provide additional information, but the policy maker would sometimes end up with a belief that leads to unnecessarily high level of uncertainty (as she could have received $\bar{\mu}$ instead). Finally if $\mu_{r_*} = \bar{\mu}$, then it is without loss to focus on a binary investigation as r_* and r_1 are interchangeable.

However, the policy maker could potentially gain when information is confidential. Recall that the policy maker gains most when the lobbyist targets her most sceptical type r_0. In addition, she gains more when her sceptical belief μ^{r_0} is lower, as the lobbyist needs to provide more evidence. However, if that belief is too low ($\mu^{r_0} < m^*(\mu^{r_1})$, the policy maker is relatively more likely to be sympathetic (observe r_1) and the lobbyist prefers to target a sympathetic policy maker (r_1).

By introducing a third signal realisation r_*, the policy maker could design a preliminary investigation such that:

1. r_* is relatively unlikely, so that the lobbyist does not want to target r_*
2. The presence of r_* reduces the probability of r_1 relatively more than the probability of r_0

As a result, the new incentive constraint of the lobbyist to target the sceptical type (r_0) is now looser. This implies that the policy maker can now choose an investigation that induces a lower μ^{r_0} while still satisfying the new incentive constraint. This new investigation can therefore make the policy maker better off, by forcing the lobbyist to produce even more evidence.

Proposition 22. *When $\underline{B} < B < \overline{B}$, there exists a preliminary investigation p with three possible realisations $\{r_0, r_*, r_1\}$ which yields a higher expected utility to the policy maker* *than the optimal investigation defined in Proposition 3.*

There are two points to note regarding this alternative investigation. First, it can only improve the policy maker's expected utility when the incentive constraint is binding. In all other cases, choosing the most informative preliminary investigation is better than any other investigation, and that most informative investigation has only two signal realisations. Second, the alternative investigation also involves some distortions. In the absence of lobbyist, the policy maker would have optimally chosen the most informative preliminary investigation with only two realisations, so choosing an investigation with three signals involves some loss of informativeness.

A.3 Proofs of Propositions in the text

Proof of Lemma 1. This results follows directly from the characterisation of the optimal information structure from Kamenica & Gentzkow (2011). In particular, if μ^r is above $\frac{1}{2}$, any persuasion strategy such that $\mu^r_{s_0} \geq \frac{1}{2}$ gives the lobbyist the same expected utility of 1. If μ^r is below $\frac{1}{2}$, the optimal persuasion strategy induces beliefs on the concave closure of $V(\mu)$.

□

Proof of Proposition 1. Let $\widetilde{U}(\mu^r)$ the indirect expected utility of the policy maker given an interim belief μ^r and the lobbyist's best-response to that public belief (as described in Lemma 1):

$$\begin{aligned}\widetilde{U}\left(\mu^r\right) &= \begin{cases}\mathbb{P}_{\pi_r}(s_0|r) + \mathbb{P}_{\pi_r}(s_1|r)\left(\frac{1}{2}\right) & \text{if } \mu^r < \frac{1}{2}\\ \mu^r & \text{if } \mu^r \geq \frac{1}{2}\end{cases}\\ &= \begin{cases}1-\mu^r & \text{if } \mu^r < \frac{1}{2}\\ \mu^r & \text{if } \mu^r \geq \frac{1}{2}\end{cases}\end{aligned}$$

This indirect expected utility is weakly convex. Indeed,

If $\boldsymbol{\mu^{r_0}, \mu^{r_1} < \bar{\mu} < \frac{1}{2}}$, or $\boldsymbol{\frac{1}{2} < \underline{\mu} < \mu^{r_0}, \mu^{r_1}}$ for any feasible μ^{r_0}, μ^{r_1}, then the indirect expected utility is linear and the policy maker is indifferent between any preliminary investigations as her expected utility is either

$$U^P(\mu^{r_0}, \mu^{r_1}) = \sum_{r\in\{r_0,r_1\}} \mathbb{P}(\mu^r)\widetilde{U}\left(\mu^r\right) = \sum_{r\in\{r_0,r_1\}} \mathbb{P}(\mu^r)\left(1-\mu^r\right) = 1-\mu_0$$

or

$$U^P(\mu^{r_0}, \mu^{r_1}) = \sum_{r\in\{r_0,r_1\}} \mathbb{P}(\mu^r)\widetilde{U}\left(\mu^r\right) = \sum_{r\in\{r_0,r_1\}} \mathbb{P}(\mu^r)\mu^r = \mu_0$$

and therefore does not depend on the choice of preliminary investigation.

If $\boldsymbol{\underline{\mu} < \frac{1}{2} < \bar{\mu}}$, then her indirect expected utility is a continuous piecewise function of two linear functions. I show that this function is convex. That is, $\forall\ \mu_L, \mu_H \in [0,1]$ and for any $\lambda \in [0,1]$, we have:

$$\lambda\widetilde{U}(\mu_L) + (1-\lambda)\widetilde{U}(\mu_H) \geq \widetilde{U}\left(\lambda\mu_L + (1-\lambda)\mu_H\right)$$

Indeed,

1. If $\mu_L < \mu_H \leq \frac{1}{2}$, then for both $\mu \in \{\mu_L, \mu_H\}$, $\widetilde{U}(\mu) = 1 - \mu$ which is linear and therefore convex.

2. If $\frac{1}{2} \leq \mu_L < \mu_H$, then for both $\mu \in \{\mu_L, \mu_H\}$, $\widetilde{U}(\mu) = \mu$ which is linear and therefore convex.

3. If $\mu_L < \frac{1}{2} < \mu_H$, and $\lambda\mu_L + (1-\lambda)\mu_H < \frac{1}{2}$ then:

$$\begin{aligned}\lambda\widetilde{U}(\mu_L) + (1-\lambda)\widetilde{U}(\mu_H) &= \lambda(1-\mu_L) + (1-\lambda)\mu_H \\ &\geq \lambda(1-\mu_L) + (1-\lambda)\frac{1}{2} \\ &= \lambda(1-\mu_L) + (1-\lambda)\left[\frac{1}{2} - \frac{1}{2} + \frac{1}{2}\right] \\ &= 1 - \left[\lambda\mu_L + (1-\lambda)\frac{1}{2}\right] \\ &\geq 1 - [\lambda\mu_L + (1-\lambda)\mu_H] = \widetilde{U}(\lambda\mu_L + (1-\lambda)\mu_H)\end{aligned}$$

   Where the first and second inequality follow from the fact that $\mu_H > \frac{1}{2}$. Similarly, if $\lambda\mu_L + (1-\lambda)\mu_H < \frac{1}{2}$, then $\lambda\widetilde{U}(\mu_L) + (1-\lambda)\widetilde{U}(\mu_H) \geq \lambda\mu_L + (1-\lambda)\mu_H = \widetilde{U}(\lambda\mu_L + (1-\lambda)\mu_H)$ where the inequality follows from $1 - \mu_L > \frac{1}{2}$.

As a result, we can apply directly results from Kamenica & Gentzkow (2011) to conclude that a preliminary investigation inducing the most extreme beliefs is optimal, and it is therefore an equilibrium for the policy maker to make full use of her expertise when her information is publicly available.

□

Proof of Lemma 2. There are two cases to consider depending on the interim beliefs of the policy maker:

1. **Lobbyist needs to persuade both types:**

   Suppose that $\mu^{r_0} < \mu^{r_1} < \frac{1}{2}$. Following the intuition from Lemma 1, to minimise the probability of s_0, the lobbyist's strategy induces $\mu^r_{s_0} = 0$ for any r. Similarly, to maximise the probability of s_1, the lobbyist's strategy induces either $\mu^{r_0}_{s_1} = \frac{1}{2}$ or $\mu^{r_1}_{s_1} = \frac{1}{2}$. These observations completely determine the two persuasion strategies π_G and π_T characterised in definitions 1 and 2.

   When he chooses π_G, the lobbyist's expected utility is simply $\mathbb{P}_{\pi_G}(s_1)$ as the realisation s_1 persuades both types in this case. When he chooses π_T, his expected utility is $\mathbb{P}(r_1)\,\mathbb{P}_{\pi_T}(s_1|r_1)$ as he only persuades type r_1 following realisation s_1. The lobbyist therefore chooses signal π_G if and only if:

$$\mathbb{P}_{\pi_G}(s_1) \geq \mathbb{P}(r_1)\,\mathbb{P}_{\pi_T}(s_1|r_1) \tag{19}$$

Let $\mu_s^r(\pi)$ the posterior induced by $s \in \{s_0, s_1\}$ from signal $\pi \in \{\pi_G, \pi_T\}$, starting from interim $\mu^r \in \{\mu^{r_0}, \mu^{r_1}\}$, and $\mu_s(\pi)$ the posterior belief induced by $s \in \{s_0, s_1\}$ from signal $\pi \in \{\pi_G, \pi_T\}$, starting from the prior μ_0.

Using the Bayes plausibility constraint, inequality (19) becomes

$$\begin{aligned}
& \frac{\mu_0 - \mu_{s_0}(\pi_G)}{\mu_{s_1}(\pi_G) - \mu_{s_0}(\pi_G)} \geq \left(\frac{\mu_0 - \mu^{r_0}}{\mu^{r_1} - \mu^{r_0}}\right)\left(\frac{\mu^{r_1} - \mu^{r_1}_{s_0}(\pi_T)}{\mu^{r_1}_{s_1}(\pi_T) - \mu^{r_1}_{s_0}(\pi_T)}\right) \\
\Leftrightarrow \quad & \frac{\mu_0(1-\mu^{r_0}) + (1-\mu_0)\mu^{r_0}}{2(1-\mu^{r_0})} \geq \frac{\mu^{r_1}(\mu_0 - \mu^{r_0})}{(\mu^{r_1} - \mu^{r_0})} \\
\Leftrightarrow \quad & -(2(\mu^{r_1} - \mu_0) + 1)(\mu^{r_0})^2 + (3\mu^{r_1} - \mu_0)\mu^{r_0} - \mu_0\mu^{r_1} \geq 0
\end{aligned} \tag{20}$$

This defines a set of pairs of beliefs (μ^{r_0}, μ^{r_1}) such that the lobbyist prefers π_G to π_T. Let, $G = \{(\mu^{r_0}, \mu^{r_1}) \mid \mu^{r_0} \geq m^*(\mu^{r_1})\}$ denote this set. The boundary of the set G is therefore given by a root of the equation:

$$H(\mu^{r_0}) = -(2(\mu^{r_1} - \mu_0) + 1)(\mu^{r_0})^2 + (3\mu^{r_1} - \mu_0)\mu^{r_0} - \mu_0\mu^{r_1} = 0 \tag{21}$$

$H(x)$ is a quadratic function of x with two roots: $\underline{x}$ and $\bar{x}$. Its first derivative is first positive up to some $x^* \in [\underline{x}, \bar{x}]$, then negative. Therefore, the function $H(x)$ is negative on $x \in [0, \underline{x}]$, positive on $[\underline{x}, \bar{x}]$, and negative on $[\bar{x}, 1]$.

Define:

$$\begin{aligned}
a(\mu^{r_1}, \mu_0) &= -(2(\mu^{r_1} - \mu_0) + 1) \\
b(\mu^{r_1}, \mu_0) &= 3\mu^{r_1} - \mu_0 \\
c(\mu^{r_1}, \mu_0) &= -\mu_0\mu^{r_1}
\end{aligned}$$

Notice that fixing μ_0 and μ^{r_1}, the function $H(x)$ defined above is a quadratic function of x with two roots:

$$\underline{x} = \frac{-b(\mu^{r_1}, \mu_0) + \sqrt{(b(\mu^{r_1}, \mu_0))^2 - 4a(\mu^{r_1}, \mu_0)c(\mu^{r_1}, \mu_0)}}{2a(\mu^{r_1}, \mu_0)}$$
$$\bar{x} = \frac{-b(\mu^{r_1}, \mu_0) - \sqrt{(b(\mu^{r_1}, \mu_0))^2 - 4a(\mu^{r_1}, \mu_0)c(\mu^{r_1}, \mu_0)}}{2a(\mu^{r_1}, \mu_0)}$$

Its first derivative $H'(\mu) = b(\mu^{r_1}, \mu_0) + 2a(\mu^{r_1}, \mu_0)x$ is positive for $x \in \left[0, \frac{-b(\mu^{r_1},\mu_0)}{2a(\mu^{r_1},\mu_0)}\right]$ since $a(\mu^{r_1}, \mu_0) < 0$ and it is negative on $\left[\frac{-b(\mu^{r_1},\mu_0)}{2a(\mu^{r_1},\mu_0)}, 1\right]$. Therefore, the function $H(x)$ is negative on $x \in [0, \underline{x}]$, positive on $[\underline{x}, \bar{x}]$, and negative on $[\bar{x}, 1]$.

Let, $m^*(\mu^{r_1})$ the lowest root of equation (21),

$$m^*(\mu^{r_1}) = \frac{3\mu^{r_1} - \mu_0 - \left[(\mu^{r_1} - \mu_0)(9\mu^{r_1} - 8\mu_0\mu^{r_1} - \mu_0\right]^{\frac{1}{2}}}{4(\mu^{r_1} - \mu_0) + 2}$$

Then we have that the lobbyist prefers π_G only if $\mu^{r_0} \geq m^*(\mu^{r_1})$.

Finally, note that for any $\mu_0 \in [0, 1]$ and $\mu^{r_1} \in [\mu_0, 1]$:

$$m^*(\mu^{r_1}) = \frac{-b(\mu^{r_1}, \mu_0) + \sqrt{(b(\mu^{r_1}, \mu_0))^2 - 4a(\mu^{r_1}, \mu_0)c(\mu^{r_1}, \mu_0)}}{2a(\mu^{r_1}, \mu_0)} \leq \mu_0$$

Indeed, re-arranging gives:

$$\begin{aligned}
&\frac{-b(\mu^{r_1}, \mu_0) + \sqrt{(b(\mu^{r_1}, \mu_0))^2 - 4a(\mu^{r_1}, \mu_0)c(\mu^{r_1}, \mu_0)}}{2a(\mu^{r_1}, \mu_0)} \leq \mu_0 \\
\Leftrightarrow \quad & \sqrt{(b(\mu^{r_1}, \mu_0))^2 - 4a(\mu^{r_1}, \mu_0)c(\mu^{r_1}, \mu_0)} \geq 2\mu_0 a(\mu^{r_1}, \mu_0) + b(\mu^{r_1}, \mu_0) \\
\Leftrightarrow \quad & b(\mu^{r_1}, \mu_0))^2 - 4a(\mu^{r_1}, \mu_0)c(\mu^{r_1}, \mu_0) \\
& \qquad \geq 4\mu_0^2 a(\mu^{r_1}, \mu_0)^2 + 4\mu_0 a(\mu^{r_1}, \mu_0)b(\mu^{r_1}, \mu_0) + b(\mu^{r_1}, \mu_0)^2 \\
\Leftrightarrow \quad & 0 \geq 4a(\mu^{r_1}, \mu_0)\left[a(\mu^{r_1}, \mu_0)\mu_0^2 + b(\mu^{r_1}, \mu_0)\mu_0 + c(\mu^{r_1}, \mu_0)\right] \\
\Leftrightarrow \quad & 0 \leq a(\mu^{r_1}, \mu_0)\mu_0^2 + b(\mu^{r_1}, \mu_0)\mu_0 + c(\mu^{r_1}, \mu_0)
\end{aligned}$$

Substituting the values of $a(\mu^{r_1}, \mu_0)$, $b(\mu^{r_1}, \mu_0)$ and $c(\mu^{r_1}, \mu_0)$, we find that this holds if

$$\begin{aligned}
& 0 \leq -\left[(2(\mu^{r_1} - \mu_0) + 1)\right]\mu_0^2 + \left[3\mu^{r_1} - \mu_0\right]\mu_0 - \mu_0\mu^{r_1} \\
\Leftrightarrow \quad & 0 \leq 2\mu_0(1 - \mu_0)(\mu^{r_1} - \mu_0)
\end{aligned}$$

which is true as $\mu_{r_1} \geq \mu_0$.

In addition, $0 \leq m^*(\mu^{r_1})$. Indeed as $a(\mu^{r_1}, \mu_0) \leq 0$,

$$\begin{aligned}
&\frac{-b(\mu^{r_1}, \mu_0) + \sqrt{(b(\mu^{r_1}, \mu_0))^2 - 4a(\mu^{r_1}, \mu_0)c(\mu^{r_1}, \mu_0)}}{2a(\mu^{r_1}, \mu_0)} \geq 0 \\
\Leftrightarrow \quad & \sqrt{(b(\mu^{r_1}, \mu_0))^2 - 4a(\mu^{r_1}, \mu_0)c(\mu^{r_1}, \mu_0)} \leq b(\mu^{r_1}, \mu_0) \\
\Leftrightarrow \quad & -4a(\mu^{r_1}, \mu_0)c(\mu^{r_1}, \mu_0) \leq 0 \\
\Leftrightarrow \quad & a(\mu^{r_1}, \mu_0)c(\mu^{r_1}, \mu_0) \geq 0
\end{aligned}$$

which is true as $c(\mu^{r_1}, \mu_0) \leq 0$.

Similarly, we can show that $\bar{x} \geq \mu_0$, so that condition (20) is satisfied for all $\mu^{r_0} \in [m^*(\mu^{r_1}), \mu_0]$. As a result, we get that for any $\mu^{r_0} \in [0, \mu_0]$, $H(\mu^{r_0})$ is positive if and only if $\mu^{r_0} \geq \underline{x}$, so the lobbyist prefers π_G only if $\mu^{r_0} \geq m^*(\mu^{r_1})$.

Finally note that the left-hand side of condition (20) is increasing in μ^{r_0} while the right-hand side is decreasing in μ^{r_0}. As μ^{r_1} increases, the right-hand side shifts down. As a result, the value of μ^{r_0}, m^*, that makes the two sides equal is decreasing in μ^{r_1}. This proves that $m^*(\mu^{r_1})$ is decreasing in μ^{r_1}.

Indeed,

$$\begin{aligned}
\frac{\partial}{\partial \mu^{r_0}} \left[\frac{\mu_0(1-\mu^{r_0}) + (1-\mu_0)\mu^{r_0}}{2(1-\mu^{r_0})} \right] &= \frac{(1-\mu_0)}{2(1-\mu^{r_0})^2} > 0 \\
\frac{\partial}{\partial \mu^{r_0}} \left[\frac{\mu^{r_1}(\mu_0 - \mu^{r_0})}{(\mu^{r_1} - \mu^{r_0})} \right] &= -\frac{\mu^{r_1}(\mu^{r_1} - \mu_0)}{(\mu^{r_1} - \mu^{r_0})^2} < 0 \\
\frac{\partial}{\partial \mu^{r_1}} \left[\frac{\mu^{r_1}(\mu_0 - \mu^{r_0})}{(\mu^{r_1} - \mu^{r_0})} \right] &= -\frac{\mu^{r_0}(\mu_0 - \mu^{r_0})}{(\mu^{r_1} - \mu^{r_0})^2} < 0
\end{aligned}$$

2. **Lobbyist only needs to persuade lower type:**

Suppose now that $\mu^{r_0} < \frac{1}{2} < \mu^{r_1}$. The lobbyist's strategy should induce $\mu^{r_0}_{s_1} = \frac{1}{2}$. If it induces a belief below $\frac{1}{2}$ then the lobbyist's payoff is weakly below that of providing no information and he would prefer an uninformative strategy. If it is strictly above $\frac{1}{2}$, then the lobbyist could increase the probability of s_1 by reducing that posterior belief without reducing his payoff. Therefore the trade-off for the lobbyist is now between a strategy such that a realisation s_0 persuades no type (i.e. $\mu^{r_1}_{s_0} < \frac{1}{2}$) and one that still persuades the high type (i.e $\mu^{r_1}_{s_0} = \frac{1}{2}$). This can be achieved with the conditional distributions π_G and π_T characterised in definitions 1 and 2.

The lobbyist now chooses π_G if and only if:

$$\begin{aligned}
& \mathbb{P}_{\pi_G}(s_1) \geq \mathbb{P}_{\pi_T}(s_1) + \mathbb{P}(r_1)\,\mathbb{P}_{\pi_T}(s_0|r_1) \\
\Leftrightarrow \quad & \frac{\mu_0 - \mu_{s_0}(\pi_G)}{\mu_{s_1}(\pi_G) - \mu_{s_0}(\pi_G)} \geq \left(\frac{\mu_0 - \mu_{s_0}(\pi_T)}{\mu_{s_1}(\pi_T) - \mu_{s_0}(\pi_T)}\right) \\
& \qquad + \left(\frac{\mu_0 - \mu^{r_0}}{\mu^{r_1} - \mu^{r_0}}\right)\left(\frac{\mu^{r_1}_{s_1}(\pi_T) - \mu^{r_1}}{\mu^{r_1}_{s_1}(\pi_T) - \mu^{r_1}_{s_0}(\pi_T)}\right) \\
\Leftrightarrow \quad & \frac{(1-\mu^{r_1})(1-2\mu^{r_0})\left[-(2(\mu^{r_1}-\mu_0)+1)(\mu^{r_0})^2 + (3\mu^{r_1}-\mu_0)\mu^{r_0} - \mu_0\mu^{r_1}\right]}{(1-\mu^{r_0})(\mu^{r_1}-\mu^{r_0})^2} \geq 0
\end{aligned}$$

This holds if and only if the following inequality is satisfied:

$$-(2(\mu^{r_1}-\mu_0)+1)(\mu^{r_0})^2 + (3\mu^{r_1}-\mu_0)\mu^{r_0} - \mu_0\mu^{r_1} \geq 0$$

The left-hand side of this inequality is $H(\mu^{r_0})$ as defined in equation (21) in the previous part of the proof. The rest of the proof therefore follows the same logic as in the previous case.

□

Proof of Proposition 2. To prove this result, I first show that the policy maker's expected utility is always strictly higher when the lobbyist chooses π_G than when her information is public for a given pair of interim beliefs (μ^{r_0}, μ^{r_1}) ('if' statement). I then show that the policy maker's expected utility is the same when the lobbyist chooses π_T as when her information is public ('only if' statement).

1. **If the incentive constraint is satisfied confidentiality is strictly preferred:** If the interim beliefs induced by the policy maker's preliminary investigation are $(\mu^{r_0}, \mu^{r_1}) \in G$, then the lobbyist chooses a general persuasion strategy π_G, which yields the following expected utility for the policy maker

$$\begin{aligned}
U^G(\mu^{r_0}, \mu^{r_1}) &= \mathbb{P}(r_0)\left[\mathbb{P}_{\pi_G}(s_0|r_0) + \mathbb{P}_{\pi_G}(s_1|r_0)\,\frac{1}{2}\right] \\
&\quad + \mathbb{P}(r_1)\left[\mathbb{P}_{\pi_G}(s_0|r_1) + \mathbb{P}_{\pi_G}(s_1|r_1)\mu^{r_1}_{s_1}\right] \\
&= (1-\mu_0)\frac{(1-2\mu^{r_0})}{(1-\mu^{r_0})} + \mu_0
\end{aligned}$$

The policy maker's expected utility under transparency is:

$$
\begin{aligned}
U^P(\mu^{r_0},\mu^{r_1}) &= \begin{cases} \mathbb{P}(s_0)\cdot 1 + \mathbb{P}(s_1)\cdot\left(\frac{1}{2}\right) & \text{if } \mu^{r_1} < \frac{1}{2} \\ \mathbb{P}(r_0)\left[\mathbb{P}_{\pi_{r_0}}(s_0|r_0)\cdot 1 + \mathbb{P}_{\pi_{r_0}}(s_1|r_0)\cdot\left(\frac{1}{2}\right)\right] + \mathbb{P}(r_1)\mu^{r_1} & \text{if } \mu^{r_1} > \frac{1}{2} \end{cases} \\
&= \begin{cases} 1-\mu_0 & \text{if } \mu^{r_1} < \frac{1}{2} \\ \left(\frac{\mu^{r_1}-\mu_0}{\mu^{r_1}-\mu^{r_0}}\right)(1-2\mu^{r_0}) + \mu_0 & \text{if } \mu^{r_1} > \frac{1}{2} \end{cases}
\end{aligned} \tag{22}
$$

Therefore, π_G always makes the policy maker better-off as

$$
\begin{aligned}
U^G(\mu^{r_0},\mu^{r_1}) &= (1-\mu_0)\frac{(1-2\mu^{r_0})}{(1-\mu^{r_0})} + \mu_0 > 1-\mu_0 = U^P(\mu^{r_0},\mu^{r_1}) \text{ if } \mu^{r_1} < \frac{1}{2} \\
U^G(\mu^{r_0},\mu^{r_1}) &= (1-\mu_0)\frac{(1-2\mu^{r_0})}{(1-\mu^{r_0})} + \mu_0 \\
&> \left(\frac{\mu^{r_1}-\mu_0}{\mu^{r_1}-\mu^{r_0}}\right)(1-2\mu^{r_0}) + \mu_0 \\
&= U^P(\mu^{r_0},\mu^{r_1}) \text{ if } \mu^{r_1} > \frac{1}{2}
\end{aligned}
$$

2. **If the incentive constraint is not satisfied, confidentiality is not strictly preferred:**

   If $(\mu^{r_0},\mu^{r_1}) \notin G$ then the lobbyist chooses a targeted strategy which yields the following expected utility for the policy maker

$$
\begin{aligned}
U^T(\mu^{r_0},\mu^{r_1}) &= \begin{cases} \mathbb{P}(r_0)\left[\mathbb{P}_{\pi_T}(s_0|r_0) + \mathbb{P}_{\pi_T}(s_1|r_0)(1-\mu^{r_0}_{s_1})\right] & \\ \qquad +\mathbb{P}(r_1)\left[\mathbb{P}_{\pi_T}(s_0|r_1) + \mathbb{P}_{\pi_T}(s_1|r_1)\left(\frac{1}{2}\right)\right] & \text{if } \mu^{r_1} < \frac{1}{2} \\ \mathbb{P}(r_0)\left[\mathbb{P}_{\pi_T}(s_0|r_0)(1-\mu^{r_0}_{s_0}) + \mathbb{P}_{\pi_T}(s_1|r_0)\left(\frac{1}{2}\right)\right] & \\ \qquad +\mathbb{P}(r_1)\left[\mathbb{P}_{\pi_T}(s_0|r_1)\left(\frac{1}{2}\right) + \mathbb{P}_{\pi_T}(s_1|r_1)\mu^{r_1}_{s_1}\right] & \text{if } \mu^{r_1} < \frac{1}{2} \end{cases} \\
&= \begin{cases} 1-\mu_0 & \text{if } \mu^{r_1} < \frac{1}{2} \\ \left(\frac{\mu^{r_1}-\mu_0}{\mu^{r_1}-\mu^{r_0}}\right)(1-2\mu^{r_0}) + \mu_0 & \text{if } \mu^{r_1} > \frac{1}{2} \end{cases}
\end{aligned} \tag{23}
$$

   Therefore, for any (μ^{r_0},μ^{r_1})

$$
U^P(\mu^{r_0},\mu^{r_1}) = U^T(\mu^{r_0},\mu^{r_1})
$$

□

Proof of Proposition 3. As before, I divide the proof into two.

1. **Lobbyist needs to persuade both types:** Suppose that for any feasible pair of beliefs, $\mu^{r_0} < \mu^{r_1} < \frac{1}{2}$.

   Claim 1: When $\bar{\mu} < \frac{1}{2}$, the policy maker is always better-off when the lobbyist chooses π_G than π_T.

   Proof: Recall that in this case, the policy maker's expected utility under transparency, given in equation (22), is $U^P(\mu^{r_0}, \mu^{r_1}) = 1 - \mu_0$, which is independent of (μ^{r_0}, μ^{r_1}) and always worse than π_G. Therefore,

   $$U^G(\mu^{r_0}, \mu^{r_1}) > U^P(\mu^{r_0}, \mu^{r_1}) = U^P(\mu_0, \mu_0)$$

   Recall that U^P and U^T yield the same expected utility, independently of (μ^{r_0}, μ^{r_1}). Therefore,

   $$U^T(\mu^{r_0}, \mu^{r_1}) = 1 - \mu_0 = U^P(\mu^{r_0}, \mu^{r_1}) = U^P(\mu_0, \mu_0)$$

   Therefore, $\forall\ \mu^{r_0}, \mu^{r_1}, \mu^{r'_0}, \mu^{r'_1}$:

   $$U^G(\mu^{r_0}, \mu^{r_1}) > U^P(\mu_0, \mu_0) = U^T\left(\mu^{r'_0}, \mu^{r'_1}\right)$$

   So the policy maker always prefers to induce the lobbyist to target both of her types (choose π_G) than to target only her sympathetic type (π_T).

   Claim 2: Given that the lobbyist chooses π_G, the policy maker prefers a preliminary investigation inducing the lowest possible sceptical belief μ^{r_0}, and is indifferent between any sympathetic belief μ^{r_1} induced by her investigation.

   Proof: The function $U^G(\mu^{r_0}, \mu^{r_1}) = (1-\mu_0)\frac{1-2\mu^{r_0}}{(1-\mu^{r_0})} + \mu_0$ is independent of μ^{r_1} and decreasing in μ^{r_0}:

   $$\frac{\partial U^G(\mu^{r_0}, \mu^{r_1})}{\partial \mu^{r_0}} = -\frac{(1-\mu_0)}{(1-\mu^{r_0})^2} \leq 0$$

   Therefore, the policy maker always prefers her preliminary investigation to induce the lowest possible μ^{r_0} subject to the incentive constraint being satisfied.

   Combining claims 1 and 2 implies that the policy maker's choice of preliminary investigation is fully determined by whether the incentive constraint is binding or not. If it is not, then she chooses the most informative preliminary investigation. If it is, then she chooses the preliminary investigation that induces the lowest μ^{r_0}

subject to the incentive constraint.

Let $\underline{B}$ solve $\underline{\mu} = m^*(\bar{\mu})$, that is:

$$\frac{\mu_0}{\mu_0 + \underline{B}(1-\mu_0)} = m^* \left(\frac{\underline{B}\mu_0}{\underline{B}\mu_0 + (1-\mu_0)} \right)$$

This threshold is given by $\underline{B} = 1 + \sqrt{2}$, and the incentive constraint for the lobbyist to choose π_G binds if and only if $B \geq \underline{B}$.

Therefore,

(a) When $B < \underline{B}$, the incentive constraint does not bind ($m^*(\bar{\mu}) < \underline{\mu}$) and the optimal preliminary investigation induces $\mu^{r_0} = \underline{\mu}$. Since $U^G(\mu^{r_0}, \mu^{r_1})$ is independent of μ^{r_1} the policy maker is indifferent between any values of μ^{r_1}, so any $(\mu^{r_0}, \mu^{r_1}) \in \{\underline{\mu}\} \times \left[(m^*)^{-1}(\underline{\mu}), \bar{\mu}\right]$ can be in the support of an equilibrium preliminary investigation. **Therefore, when $\boldsymbol{B < \underline{B}}$, it is an equilibrium for the policy maker to choose the most informative preliminary investigation.**

(b) When $B > \underline{B}$, the constraint binds ($m^*(\bar{\mu}) > \underline{\mu}$) and the optimal preliminary investigation induces $\mu^{r_0} = m^*(\mu^{r_1})$. Since $m^*(\mu^{r_1})$ is decreasing in μ^{r_1} and $U^G(\mu^{r_0}, \mu^{r_1})$ is independent of μ^{r_1}, it is optimal to induce $\mu^{r_1} = \bar{\mu}$. **Therefore, when $\boldsymbol{B \geq \underline{B}}$, the only equilibrium is for the policy maker to choose a preliminary investigation that induces $\boldsymbol{\mu^{r_0} = m^*(\bar{\mu})}$ and $\boldsymbol{\mu^{r_1} = \bar{\mu}}$.**

2. **Lobbyist only needs to persuade lower type:**

Suppose now that beliefs such that $\mu^{r_0} < \frac{1}{2} < \mu^{r_1}$ are feasible. Note that $U^G(\mu^{r_0}, \mu^{r_1})$ is unchanged in this case. I show that the policy maker now sometimes prefer to induce π_T.

Claim 3: If the lobbyist were to choose π_T, the policy maker would choose the most informative investigation $\bar{p}$.

Proof: Recall that the policy maker's indirect utility as a function of (μ^{r_0}, μ^{r_1}) in this case is:

$$U^T(\mu^{r_0}, \mu^{r_1}) = \frac{(\mu^{r_1} - \mu_0)(1-\mu^{r_0}) + (\mu_0 - \mu^{r_0})\mu^{r_1}}{\mu^{r_1} - \mu^{r_0}} \tag{24}$$

Therefore, $U^T(\mu^{r_0},\mu^{r_1})$ is increasing in μ^{r_1} and decreasing in μ^{r_0}:

$$\frac{\partial U^T(\mu^{r_0},\mu^{r_1})}{\partial \mu^{r_0}} = -\frac{(2\mu^{r_1}-1)(\mu^{r_1}-\mu_0)}{(\mu^{r_1}-\mu^{r_0})^2} \leq 0$$
$$\frac{\partial U^T(\mu^{r_0},\mu^{r_1})}{\partial \mu^{r_1}} = \frac{(1-2\mu^{r_0})(\mu_0-\mu^{r_0})}{(\mu^{r_1}-\mu^{r_0})^2} \geq 0$$

As a result, the optimal preliminary investigation given π_T induces interim beliefs $(\underline{\mu},\bar{\mu})$.

Claim 4: If $B < \underline{B}$, the policy maker strictly prefers the most informative preliminary investigation.

Proof: As shown above, if $B < \underline{B}$, the policy maker induces the lobbyist to choose π_G over π_T even when using the most informative investigation. From the proof of Proposition 2, we know that $U^G(\mu^{r_0},\mu^{r_1}) > U^T(\mu^{r_0},\mu^{r_1})$. Therefore $U^G\left(\underline{\mu},\bar{\mu}\right) > U^T\left(\underline{\mu},\bar{\mu}\right)$. From claim 3, we know that $U^T\left(\underline{\mu},\bar{\mu}\right) \geq U^T(\mu^{r_0},\mu^{r_1})$ for any feasible (μ^{r_0},μ^{r_1}).

Therefore,

$$U^G\left(\underline{\mu},\bar{\mu}\right) > U^T\left(\underline{\mu},\underline{\mu}\right) \geq U^T(\mu^{r_0},\mu^{r_1})$$

Finally, from claim 2, we have $U^G\left(\underline{\mu},\bar{\mu}\right) \geq U^G(\mu^{r_0},\mu^{r_1})$. Therefore if the incentive constraint does not bind ($B < \underline{B}$), the policy maker prefers the most informative preliminary information $\bar{p}$.

Claim 5: If $B > \underline{B}$, there exists $\overline{B} > \underline{B}$ such that the policy maker prefers a preliminary investigation such that $\mu^{r_0} = m^*(\bar{\mu})$ and to induce the lobbyist to choose π_G if $B \leq \overline{B}$, and prefers a preliminary investigation such that $\mu^{r_0} = \underline{\mu}$ and to induce the lobbyist to choose π_T if $B \geq \overline{B}$.

Proof: First recall that if $B > \underline{B}$, the incentive constraint binds so given claim 2, the optimal preliminary investigation that induces π_G generates interim beliefs $\mu^{r_0} = m^*(\bar{\mu})$ and $\mu^{r_1} = \bar{\mu}$. Given claim 3, if the policy maker's investigation induces π_T, then it is optimal to generate interim beliefs $\mu^{r_0} = \underline{\mu}$ and $\mu^{r_1} = \bar{\mu}$. To show the existence and uniqueness of $\overline{B}$, I proceed in three steps.

Step 1: At $B = \underline{B}$, $m^*(\bar{\mu}) = \underline{\mu}$ so $U^G(m^*(\bar{\mu}),\bar{\mu}) = U^G\left(\underline{\mu},\bar{\mu}\right) > U^T\left(\underline{\mu},\bar{\mu}\right)$ by claim 4.

Step 2: $U^G(m^*(\bar{\mu}),\underline{\mu}) - U^T(\underline{\mu},\bar{\mu})$ is strictly decreasing in B for $B > \underline{B}$. Indeed, we can write

$$U^G\left(m^*(\bar{\mu}),\underline{\mu}\right)-U^T\left(\underline{\mu},\bar{\mu}\right)=\frac{1}{B+1}-\frac{2B}{3B+\sqrt{9B^2-10B+1}-1}\mu_0$$

And taking derivatives with respect to B proves the result:

$$\begin{aligned}&\frac{\partial\left[U^G\left(m^*(\bar{\mu}),\underline{\mu}\right)-U^T\left(\underline{\mu},\bar{\mu}\right)\right]}{\partial B}\\&\qquad\qquad=-\frac{1}{(B+1)^2}-\mu_0\frac{-10B+2-2\sqrt{9B^2-10B+1}}{(3B+\sqrt{9B^2-10B+1}-1)^2\sqrt{9B^2-10B+1}}\end{aligned}$$

Therefore, as $-10B+2-2\sqrt{9B^2-10B+1}<0$, $U^G(m^*(\bar{\mu}),\underline{\mu})-U^T(\underline{\mu},\bar{\mu})$ is strictly decreasing if and only if:

$$\mu_0<\frac{(3B+\sqrt{9B^2-10B+1}-1)^2\sqrt{9B^2-10B+1}}{(B+1)^2(-2+10B+2\sqrt{9B^2-10B+1})} \tag{25}$$

Second, the right-hand side of (25), $\frac{(3B+\sqrt{9B^2-10B+1}-1)^2\sqrt{9B^2-10B+1}}{(B+1)^2(-2+10B+2\sqrt{9B^2-10B+1})}$ is increasing in B:

$$\begin{aligned}&\frac{\partial}{\partial B}\left[\frac{(3B+\sqrt{9B^2-10B+1}-1)^2\sqrt{9B^2-10B+1}}{(B+1)^2(-2+10B+2\sqrt{9B^2-10B+1})}\right]=\\&\quad=\frac{1}{(B+1)^3(5B-1+\sqrt{9B^2-10B+1})^2\sqrt{9B^2-10B+1}}\\&\qquad\times 2(3B+\sqrt{9B^2-10B+1}-1)\\&\qquad\quad\times(165B^3-207B^2+63B-5+(57B^2-38B+5)\sqrt{9B^2-10B+1})\end{aligned}$$

This is positive if and only if:

$$165B^3-207B^2+63B-5+(57B^2-38B+5)\sqrt{9B^2-10B+1}\geq 0$$

This always holds for $B>1$ as $57B^2-38B+5>0$ if and only if $B>\frac{1}{3}+\frac{2\sqrt{19}}{57}\approx 0.49$, while $165B^3-207B^2+63B-5$ has three real roots all smaller than 1 and is greater than 0 for B greater than 1.

Finally, if the right-hand side of (25) is increasing in B, and at $B=1+\sqrt{2}$ we have

$$\frac{(3B+\sqrt{9B^2-10B+1}-1)^2\sqrt{9B^2-10B+1}}{(B+1)^2(-2+10B+2\sqrt{9B^2-10B+1})}=\frac{(3+2\sqrt{2})^2(4+\sqrt{2})}{(2+\sqrt{2})^2(4+3\sqrt{2})}>1>\mu_0$$

Then for any $B > \underline{B} = 1 + \sqrt{2}$, $U^G(m^*(\bar{\mu}), \underline{\mu}) - U^T(\underline{\mu}, \bar{\mu})$ is decreasing in B.

Step 3: As $B \to +\infty$, $U^G(m^*(\bar{\mu}), \underline{\mu}) < U^T(\underline{\mu}, \bar{\mu})$.

We know that $\lim_{B \to +\infty} (\underline{\mu}, \bar{\mu}) = (0, 1)$, and $\lim_{B \to +\infty} m^*(\bar{\mu}) = m^*(1) = \frac{\mu_0}{1-2\mu_0}$ and since $U^G(\mu^{r_0}, \mu^{r_1})$ is continuous in μ^{r_0}, we have:

$$\lim_{B \to +\infty} U^G(m^*(\bar{\mu}), \bar{\mu}) = \lim_{B \to +\infty} (1 - \mu_0) \frac{1 - 2m^*(\bar{\mu})}{(1 - m^*(\bar{\mu}))} + \mu_0 = 1 - \frac{\mu_0}{3}$$

In addition, $\lim_{B \to +\infty} U^T(\underline{\mu}, \bar{\mu}) = (1 - \mu_0) + \mu_0 = 1$. Therefore, we have:

$$\lim_{B \to +\infty} U^T(\underline{\mu}, \bar{\mu}) = 1 > 1 - \frac{\mu_0}{3} = \lim_{B \to +\infty} U^G(m^*(\bar{\mu}), \bar{\mu})$$

Therefore, $\lim_{B \to +\infty} U^T(\underline{\mu}, \bar{\mu}) > \lim_{B \to +\infty} U^G(m^*(\bar{\mu}), \bar{\mu})$ and the policy maker eventually prefers to give up on trying to induce π_G.

Combining steps 1, 2 and 3 and the intermediate value theorem implies that there exists a unique $\overline{B} > \underline{B}$ that satisfies claim 5.

Claims 1 to 5 then imply that the policy maker makes full use of her expertise if either $B < \underline{B}$ or $B > \overline{B}$ and otherwise distorts her investigation such that $\mu^{r_0} = m^*(\bar{\mu})$.

□

Proof of Proposition 4. To prove Proposition 4, I first derive the equilibrium value of confidentiality for all possible values of B and μ_0 and then take derivatives with respect to B and μ_0 for each possible case.

Let B_H the value of B such that, given some prior $\mu_0 < \frac{1}{2}$, $\bar{\mu} = \frac{1}{2}$. Similarly, let m_H the value of μ_0 such that, given some expertise B, $\bar{\mu} = \frac{1}{2}$. It is easy to verify that $B_H(\mu_0) = \frac{1-\mu_0}{\mu_0}$ and $m_H(B) = \frac{1}{B+1}$.

Let $\underline{U}^P(\underline{\mu}, \bar{\mu}) = 1 - \mu_0$ and $\overline{U}^P(\underline{\mu}, \bar{\mu}) = \left(\frac{\bar{\mu} - \mu_0}{\bar{\mu} - \underline{\mu}}\right)(1 - 2\underline{\mu}) + \mu_0$ the possible forms of the policy maker's equilibrium expected utility under transparency, as per equation (22).

The value of confidentiality is then

$$W(B, \mu_0) = \begin{cases} U^G(\underline{\mu}, \bar{\mu}) - \underline{U}^P(\underline{\mu}, \bar{\mu}) & \text{if } B < \min\{\underline{B}, B_H\} \\ U^G(\underline{\mu}, \bar{\mu}) - \overline{U}^P(\underline{\mu}, \bar{\mu}) & \text{if } B \in (B_H, \underline{B}) \\ U^G(m^*(\bar{\mu}), \bar{\mu}) - \underline{U}^P(\underline{\mu}, \bar{\mu}) & \text{if } B \in (\underline{B}, B_H) \\ U^G(m^*(\bar{\mu}), \bar{\mu}) - \overline{U}^P(\underline{\mu}, \bar{\mu}) & \text{if } B > \max\{\underline{B}, B_H\} \end{cases}$$

For each case, we can simplify the expressions and take partial derivatives.

1. **Case 1:** If $B < \min\{\underline{B}, B_H\}$,

   $W_1(B, \mu_0) = \left[(1-\mu_0)\frac{1-2\underline{\mu}}{1-\underline{\mu}} + \mu_0\right] - [1-\mu_0] = \frac{\mu_0(B-1)}{B}$.

   Therefore, $\frac{\partial W_1(B,\mu_0)}{\partial B} = \frac{\mu_0}{B^2} > 0$ and $\frac{\partial W_1(B,\mu_0)}{\partial \mu_0} = \frac{(B-1)}{B} > 0$.

2. **Case 2:** If $B \in (B_H, \underline{B})$,

   $W_2(B, \mu_0) = \left[(1-\mu_0)\frac{1-2\underline{\mu}}{1-\underline{\mu}} + \mu_0\right] - \left[\frac{(\bar{\mu}-\mu_0)}{\bar{\mu}-\underline{\mu}}(1-2\underline{\mu}) + \mu_0\right] = \frac{(1-\mu_0)B-\mu_0}{B(B+1)}$

   (a) For B: $\frac{\partial W_2(B,\mu_0)}{\partial B} = \frac{\mu_0(2B+1)-B^2(1-\mu_0)}{(B(B+1))^2} \leq 0$ if $B \geq \frac{\mu_0+\sqrt{\mu_0}}{(1-\mu_0)}$. This follows from the fact that $\frac{\mu_0(2B+1)-B^2(1-\mu_0)}{(B(B+1))^2} \leq 0$ if and only if $\mu_0(2B+1) - B^2(1-\mu_0) > 0$. The left-hand side is a quadratic equation in B with a negative coefficient on the squared term and with positive root $B^* = \frac{\mu_0+\sqrt{\mu_0}}{(1-\mu_0)}$. Note that B^* can be less than 1 for μ_0 sufficiently low in which case the value of confidentiality is always decreasing in this range.

   (b) For μ_0, $\frac{\partial W_2(B,\mu_0)}{\partial \mu_0} = -\frac{1}{B} < 0$.

3. **Case 3:** If $B \in (\underline{B}, B_H)$,

$$W_3(B, \mu_0) = \left[\frac{(1-\mu_0)(1-2m^*(\bar{\mu}))}{(1-m^*(\bar{\mu}))} + \mu_0\right] - [1-\mu_0]$$
$$= \frac{\mu_0\left(4\mu_0 B - 3B + \sqrt{9B^2-10B+1} - 4\mu_0 + 3\right)}{\mu_0\sqrt{9B^2-10B+1} + 2 + 3(B-1)\mu_0}$$

   (a) For B: $\frac{\partial W_3(B,\mu_0)}{\partial B} = -(1-\mu_0)\frac{1}{(1-m^*(\bar{\mu}))^2}\frac{\partial m^*(\bar{\mu})}{\partial \bar{\mu}}\frac{\partial \bar{\mu}}{\partial B} > 0$ as $\frac{\partial m^*(\bar{\mu})}{\partial \bar{\mu}} < 0$ and $\frac{\partial \bar{\mu}}{\partial B} > 0$.

   (b) For μ_0: note that we can simplify $W_3(B, \mu_0)$ to express it as a linear function of μ_0:

$$W_3(B, \mu_0) = \left(\frac{B + \sqrt{9B^2-10B+1} - 1}{3B + \sqrt{9B^2-10B+1} - 1}\right)\mu_0$$

   Therefore,

$$\frac{\partial W_3(B, \mu_0)}{\partial \mu_0} = \frac{B + \sqrt{9B^2-10B+1} - 1}{3B + \sqrt{9B^2-10B+1} - 1}$$

   Which is greater than zero as $B > 1$, $3B > 1$ and $\sqrt{9B^2-10B+1} > 0$.

4. **Case 4:** If $B > \max\{\underline{B}, B_H\}$,

$$\begin{aligned}W_4(B,\mu_0) &= \frac{(1-\mu_0)(1-2m^*(\bar{\mu}))}{(1-m^*(\bar{\mu}))} - \frac{B(1-\mu_0)-\mu_0(1-q)}{q(B+1)} \\ &= \frac{1}{(B+1)(3B\mu_0+\sqrt{9B^2-10B+1}\mu_0-3\mu_0+2)} \\ &\quad \times \Big[(B+2-\mu_0(B+1))\mu_0\sqrt{9B^2-10B+1}+2+(B^2-1)\mu_0^2 \\ &\quad + (-3B^2+B-2)\mu_0)\Big]\end{aligned}$$

Note first that $W_4(B,\mu_0)$ can be re-written as a linear function of μ_0:

$$W_4(B,\mu_0) = \frac{1}{B+1} - \left(\frac{2B}{3B+\sqrt{9B^2-10B+1}-1}\right)\mu_0$$

(a) For B: note then that the derivative with respect to B is:

$$\frac{\partial W_4(B,\mu_0)}{\partial B} = -\frac{1}{(B+1)^2} + \frac{10B-2+2\sqrt{9B^2-10B+1}}{(3B+\sqrt{9B^2-10B+1}-1)^2\sqrt{9B^2-10B+1}}\mu_0$$

Since $10B-2+2\sqrt{9B^2-10B+1} > 0$, this is negative if and only if:

$$\mu_0 < \frac{(3B+\sqrt{9B^2-10B+1}-1)^2\sqrt{9B^2-10B+1}}{(B+1)^2(10B-2+2\sqrt{9B^2-10B+1})} \tag{26}$$

Note that this is the same condition as (25).

(b) For μ_0: we simply need to note that $\frac{10B-2+2\sqrt{9B^2-10B+1}}{(3B+\sqrt{9B^2-10B+1}-1)^2\sqrt{9B^2-10B+1}} > 0$ since $10B > 10 > 2$ and all other terms are squared or square roots and therefore positive, so that

$$\frac{\partial W_4(B,\mu_0)}{\partial \mu_0} = -\frac{10B-2+2\sqrt{9B^2-10B+1}}{(3B+\sqrt{9B^2-10B+1}-1)^2\sqrt{9B^2-10B+1}} < 0$$

As a result, we can conclude that

1. When $B < \min\{\underline{B}, B_H\}$, $W(B,\mu_0)$ increases in B.
2. When $B \in (B_H, \underline{B},)$, $W(B,\mu_0)$ increases in B if $B \le \frac{\mu_0+\sqrt{\mu_0}}{(1-\mu_0)}$ and decreases otherwise.
3. When $B \in (\underline{B}, B_H)$, $W(B,\mu_0)$ increases in B.
4. When $B > \max\{\underline{B}, B_H\}$, $W(B,\mu_0)$ decreases in B.

And that $W(B, \mu_0)$ increases in μ_0 if $\mu_0 < m_H$, but decreases in μ_0 if $\mu_0 > m_H$.

Finally, when $B > \overline{B}$, the value of confidential information is 0 as the policy maker prefers to let the lobbyist choose π_T and make full use of her expertise, which yields the same expected utility as transparency.

$\square$

Proof of Proposition 5. Note that the set of equilibria under public information is the same as that characterised by Lemma 1 and Proposition 1 since when the design of the preliminary investigation is public, the policy maker can commit to it. Therefore, the equilibrium expected utility from a public preliminary investigation is given by equation (22).

Under confidential information, there can be multiple equilibria. For any equilibrium preliminary investigation p, the lobbyist's best response is determined by Lemma 2.

However, the policy maker's best response is different than in the commitment case since deviating to a different p no longer affects the strategy chosen by the lobbyist, as that deviation is not observable. I prove the proposition using the following two lemmas.

Lemma 5. *In any equilibrium in which $\pi_{\mathbf{G}}$ is played, we must have $\mu^{r_0} = \underline{\mu}$. Such an equilibrium exists if and only if $B \leq \underline{B}$.*

Proof. The proof proceeds in three steps.

Claim 1: If an equilibrium exists where $\pi_{\mathbf{G}}$ is played, then $\mu^{r_0} = \underline{\mu}$.

Proof: Suppose to the contrary that there was an equilibrium with $\mu^{r_0} > \underline{\mu}$ and a corresponding π_G. Then the policy maker's expected utility from choosing a p' inducing beliefs $(\mu^{r_0'}, \mu^{r_1})$ such that $\mu^{r_0'} \leq \mu^{r_0}$ is:

$$\begin{aligned} U^G(\mu^{r_0'}, \mu^{r_1}) = \mathbb{P}_{\pi_G}(s_0) &+ \left(\frac{\mu^{r_1} - \mu_0}{\mu^{r_1} - \mu^{r_0'}}\right)\left[\mu^{r_0'} + (1-\mu^{r_0'})\frac{\mu^{r_0}}{(1-\mu^{r_0})}\right] \\ &\times \left(1 - \frac{\mu^{r_0'}(1-\mu^{r_0})}{\mu^{r_0'}(1-\mu^{r_0}) + (1-\mu^{r_0'})\mu^{r_0}}\right) \\ &+ \left(\frac{\mu_0 - \mu^{r_0'}}{\mu^{r_1} - \mu^{r_0'}}\right)\left[\mu^{r_1} + (1-\mu^{r_1})\frac{\mu^{r_0}}{(1-\mu^{r_0})}\right] \\ &\times \left(\frac{\mu^{r_1}(1-\mu^{r_0})}{\mu^{r_1}(1-\mu^{r_0}) + (1-\mu^{r_1})\mu^{r_0}}\right) \end{aligned}$$

A deviation to $\mu^{r_0'} \leq \mu^{r_0}$ does not affect the strategy of the lobbyist, but changes (1) the relative likelihood of the two types of the policy maker and (2) the expected payoff conditional on r_0 and s_1, which now becomes $1 - \mu_{s_1}^{r_0'}$ instead of $\frac{1}{2}$. The derivative of this expected utility function with respect to $\mu^{r_0'}$ is negative, and therefore, it is always profitable to deviate to some $\mu^{r_0'} < \mu^{r_0}$.

Indeed,

$$\frac{\partial U^G(\mu^{r_0'}, \mu^{r_1})}{\partial \mu^{r_0'}} = -\frac{(\mu^{r_1} - \mu^{r_0})(\mu^{r_1} - \mu_0)}{(1 - \mu^{r_0})(\mu^{r_1} - \mu^{r_0'})^2} \leq 0$$

<u>Claim 2:</u> If the policy maker chooses a preliminary investigation p that induces interim beliefs $(\underline{\mu}, \mu^{r_1})$ and a corresponding persuasion strategy π_G, the policy maker has no incentives to deviate to p' such that $\mu^{r_0'} \geq \underline{\mu}$.

<u>Proof:</u> Such a deviation gives expected utility $U^G(\mu^{r_0'}, \mu^{r_1}) = \mathbb{P}_{\pi_G}(s_0) + \mu_0$. Deviating to $\mu^{r_0'} \geq \underline{\mu}$ changes the expected payoff conditional on r_0 and s_1, which now becomes $\mu_{s_1}^{r_0'} \geq \frac{1}{2}$ instead of $\frac{1}{2}$. Because the indirect expected utility is linear in this region, the deviation payoff is independent of $\mu^{r_0'}$ and there can be no gain.

<u>Claim 3:</u> If the policy maker chooses a preliminary investigation p that induces interim beliefs $(\underline{\mu}, \mu^{r_1})$ and a corresponding persuasion strategy π_G, the policy maker has no incentives to deviate to p' such that $\mu^{r_1'} \neq \mu^{r_1}$.

<u>Proof:</u> Such a deviation gives an expected payoff of $U^G(\mu^{r_0}, \mu^{r_1'}) = \mathbb{P}_{\pi_G}(s_0) + \mu_0$. Deviating to $\mu^{r_1'} \neq \mu^{r_1}$ changes the expected payoff conditional on r_1 and s_1, which now becomes $\mu_{s_1}^{r_1'} \neq \mu_{s_1}^{r_1} \geq \frac{1}{2}$, but such that $\mu_{s_1}^{r_1'} \geq \frac{1}{2}$. Because the indirect expected utility is linear in this region, the deviation payoff is independent of μ^{r_1} and there can be no gain.

Combining claims 2 and 3 implies that, given some lobbyist strategy π_G corresponding to some beliefs $(\mu^{r_0}, \mu^{r_1}) = (\underline{\mu}, \mu^{r_1})$, the policy maker does not deviate to any other information structure. Given Lemma 2, if $(\underline{\mu}, \mu^{r_1}) \in G$, that is, if $B \leq \underline{B}$ (and μ^{r_1} not too small), the lobbyist does want to play π_G so this is an equilibrium. **This proves the 'if' part of the existence statement.**

Finally, given Lemma 2, if $B > \underline{B}$, then $(\underline{\mu}, \mu^{r_1}) \notin G$ for any $\mu^{r_1} \in [\mu_0, \bar{\mu}]$, so claim 1 implies that the policy maker always deviates from an investigation inducing $(\mu^{r_0}, \mu^{r_1}) \in G$ to some $\mu^{r_0'} < \mu^{r_0}$. Therefore if $B > \underline{B}$ there does not exist an equilibrium in which π_G is played. **This proves the 'only if' part of the existence statement.**

□

Lemma 6. *If $B > \underline{B}$, then an equilibrium always exists, the lobbyists plays strategy π_T and the policy maker's payoff is the same as if information was public.*

Proof. If $B > \underline{B}$, then Lemma 5 implies that the only possible equilibrium involves the lobbyist choosing π_T. I show that there is indeed a choice of p from which the policy maker does not deviate.

Claim 1: If $\bar{\mu} \leq \frac{1}{2}$ $(B \leq B_H)$ the policy maker does not deviate from any preliminary investigation inducing $\mu^{r_1} = \bar{\mu}$, and any μ^{r_0} such that $(\mu^{r_0}, \bar{\mu}) \notin G$.

Proof: The policy maker's expected utility in equilibrium is $U^T(\mu^{r_0}, \mu^{r_1})$ as defined in equation (23). Deviating to $\mu^{r'_1} < \bar{\mu}$ gives expected utility of $U^T(\mu^{r_0}, \mu^{r'_1}) = \mathbb{P}_{\pi_T}(s_0) + \frac{(1-\mu_0)\bar{\mu}}{1-\bar{\mu}}$. This is independent of $\mu^{r'_1}$, so the policy maker would not deviate.

Secondly, deviating to $\mu^{r'_0} \neq \mu^{r_0}$ gives expected utility of $U^T(\mu^{r'_0}, \bar{\mu}) = \mathbb{P}_{\pi_T}(s_0) + \frac{(1-\mu_0)\bar{\mu}}{1-\bar{\mu}}$. This is independent of $\mu^{r'_0}$, so the policy maker would not deviate.

Claim 2: If $\bar{\mu} > \frac{1}{2}$ $(B > B_H)$ a preliminary investigation inducing $\mu^{r_0} = \underline{\mu}$ and $\mu^{r_1} = \bar{\mu}$ and a persuasion strategy π_T is the only equilibrium strategy profile.

Proof: The policy maker's expected utility in equilibrium is $U^T(\underline{\mu}, \bar{\mu})$ as defined in equation (23).

Deviating to an investigation that induces a pair of interim beliefs $(\mu^{r'_0}, \mu^{r'_1})$ such that $\mu^{r'_0} \geq \underline{\mu}$ and $\mu^{r'_1} \leq \bar{\mu}$ gives:

$$U^T(\mu^{r'_0}, \mu^{r'_1}) = \frac{\bar{\mu}(1 + \mu_0 - 2\underline{\mu}) - \mu_0(1 - \underline{\mu})}{\bar{\mu} - \underline{\mu}}$$

Which is the same as when $\mu^{r'_0} = \underline{\mu}$ and $\mu^{r'_1} = \bar{\mu}$, so the policy maker does not deviate to any $\mu^{r'_0} > \underline{\mu}$ and $\mu^{r'_1} < \bar{\mu}$.

Finally, we can show that this is the only equilibrium. Suppose there was an equilibrium such that interim beliefs were (μ^{r_0}, μ^{r_1}) such that $\mu^{r_0} \geq \underline{\mu}$ and $\mu^{r_1} \leq \bar{\mu}$ with at least one inequality strict. By evaluating the derivative of each payoff function when deviating, we can show that:

1. If $\mu^{r_0} > \underline{\mu}$, the policy maker would always deviate to inducing $\mu^{r'_0} < \mu^{r_0}$.

2. If $\mu^{r_1} < \bar{\mu}$, the policy maker would always deviate to inducing $\mu^{r'_1} > \mu^{r_1}$.

Therefore, claim 1 implies that there exists an equilibrium in which π_T is played when $\bar{\mu} < \frac{1}{2}$ and claim 2 implies that there exists a unique equilibrium in which π_T is played when $\bar{\mu} > \frac{1}{2}$, and that this equilibrium involves $\mu^{r_0} = \underline{\mu}$ and $\mu^{r_1} = \bar{\mu}$.

Finally, since the policy maker's expected utility is the same under confidentiality with π_T and transparency (Proposition 2), the two regimes yield the same utility whenever $B > \underline{B}$. □

Lemmas 5 and 6 imply

1. If $B < \underline{B}$, then $\bar{p}$ and π_G is an equilibrium strategy profile.

2. If $\underline{B} < B$, then $\bar{p}$ and π_G is no longer an equilibrium strategy profile, so any equilibrium under confidential information induces π_T which gives the same expected utility to the policy maker as transparency.

This concludes the proof of Proposition 5.

□

Proof of Proposition 6. The influence function depends on whether $B > \underline{B}$ and $\bar{\mu} < \frac{1}{2}$. Substituting beliefs using the Bayes plausibility constraints in expression (1.4) and re-arranging gives

$$F(B,\mu_0) = \begin{cases} \frac{(B+1)\mu_0}{B} & \text{if } B < \min\{\underline{B}, B_H\} \\ \mu_0 + (1-\mu_0)\frac{m^*(\bar{\mu})}{(1-m^*(\bar{\mu}))} & \text{if } B \in (\underline{B}, B_H) \\ \frac{(3\mu_0-1)B+\mu_0}{B(B+1)} & \text{if } B \in (B_H, \underline{B}) \\ \mu_0 + (1-\mu_0)\frac{m^*(\bar{\mu})}{(1-m^*(\bar{\mu}))} - \frac{\mu_0-\underline{\mu}}{\bar{\mu}-\underline{\mu}} & \text{if } \max\{\underline{B}, B_H\} < B < \overline{B} \\ \frac{2\mu_0}{B+1} & \text{if } \overline{B} < B \end{cases}$$

For each case, I take the partial derivatives with respect to B and μ_0.

1. **Case 1:** If if $B < \min\{\underline{B}, B_H\}$, $\frac{\partial F(B,\mu_0)}{\partial B} = -\frac{\mu_0}{B^2} < 0$ and $\frac{\partial F(B,\mu_0)}{\partial \mu_0} = \frac{B+1}{B} > 0$.

2. **Case 2:** We know that $m^*(\bar{\mu})$ is decreasing in B (as $m^*(\mu^{r_1})$ is decreasing in μ^{r_1} and $\bar{\mu}$ is increasing in B). In addition, we know that $\frac{m}{1-m}$ is increasing in m. Therefore, $F_2(B,\mu_0) = \mu_0 + (1-\mu_0)\frac{m^*(\bar{\mu})}{(1-m^*(\bar{\mu}))}$ is decreasing in B.

   Secondly, I show that $(1-\mu_0)\frac{m^*(\bar{\mu})}{(1-m^*(\bar{\mu}))}$ is increasing in μ_0. Note that $F_2(B,\mu_0)$ can be re-written as a linear function of μ_0:

$$\begin{aligned} \mu_0 + (1-\mu_0)\frac{m^*(\bar{\mu})}{(1-m^*(\bar{\mu}))} &= \\ &= \frac{\mu_0\left(2\mu_0(B-1) + \mu_0\sqrt{9B^2-10B+1} + 3B + 1 - \sqrt{9B^2-10B+1}\right)}{\mu_0\sqrt{9B^2-10B+1} + 3(B-1)\mu_0 + 2} \\ &= \left(\frac{5B-1+\sqrt{9B^2-10B+1}}{3B-1+\sqrt{9B^2-10B+1}}\right)\mu_0 \end{aligned}$$

   Since $\frac{5B-1+\sqrt{9B^2-10B+1}}{3B-1+\sqrt{9B^2-10B+1}} > 0$, $F_2(B,\mu_0)$ is an increasing function of μ_0.

3. **Case 3:** Since $F_3(B,\mu_0) = \frac{(3\mu_0-1)B+\mu_0}{B(B+1)}$, $\frac{\partial F_3(B,\mu_0)}{\partial B} = \frac{(1-3\mu_0)B^2-2\mu_0 B-\mu_0}{B^2(B+1)^2}$. This is

decreasing if and only if:

$$\begin{aligned}(1-3\mu_0)B^2-2\mu_0 B-\mu_0\leq 0 \quad &\Leftrightarrow \quad \frac{1-3\mu_0}{\mu_0}\leq\frac{2B+1}{B^2}\\ &\Leftrightarrow \quad \frac{1-3\mu_0}{\mu_0}+1\leq\frac{2B+1}{B^2}+1\\ &\Leftrightarrow \quad \frac{1-\mu_0}{\mu_0}\leq\left(\frac{B+1}{B}\right)^2+1\end{aligned}$$

In addition, since in case 3, $B_H < B < \underline{B}$, and since $\frac{B+1}{B}$ is decreasing in B, $\left(\frac{\underline{B}+1}{\underline{B}}\right)^2 < \left(\frac{B+1}{B}\right)^2$, and as $\underline{B} = 1+\sqrt{2} < \left(\frac{2+\sqrt{2}}{1+\sqrt{2}}\right)^2 + 1 = \left(\frac{\underline{B}+1}{\underline{B}}\right)^2 + 1$, we have:

$$\frac{(1-\mu_0)}{\mu_0} = B_H < \underline{B} < \left(\frac{\underline{B}+1}{\underline{B}}\right)^2 + 1 < \left(\frac{B+1}{B}\right)^2 + 1$$

And therefore $F_3(B,\mu_0)$ is decreasing in B.

Secondly, note that $\frac{\partial F_3(B,\mu_0)}{\partial\mu_0} = \frac{3B+1}{B(B+1)} > 0$, so $F_3(B,\mu_0)$ is increasing in μ_0.

4. **Case 4:**

   (a) **For $\boldsymbol{B}$:** Recall that $F_4(B,\mu_0) = \mu_0 + (1-\mu_0)\frac{m^*(\bar{\mu})}{(1-m^*(\bar{\mu}))} - \frac{\mu_0-\underline{\mu}}{\bar{\mu}-\underline{\mu}}$. The derivative with respect to B is therefore:

$$\frac{\partial F_4(B,\mu_0)}{\partial B} = \frac{\partial(1-\mu_0)\frac{m^*(\bar{\mu})}{(1-m^*(\bar{\mu}))}}{\partial B} - \frac{2\mu_0-1}{(B+1)^2}$$

   It is possible for $F_4(B,\mu_0)$ to be increasing or decreasing in B when $\mu_0 < \frac{1}{2}$. Indeed, it is increasing if:

$$\frac{1-2\mu_0}{(B+1)^2} > -\left[\frac{\partial(1-\mu_0)\frac{m^*(\bar{\mu})}{(1-m^*(\bar{\mu}))}}{\partial B}\right]$$

   This can be re-written as:

$$\frac{1-2\mu_0}{(B+1)^2} > \left(\frac{10B-2+2\sqrt{9B^2-10B+1}}{\sqrt{9B^2-10B+1}(3B+\sqrt{9B^2-10B+1})^2}\right)\mu_0$$

So $F_4(B,\mu_0)$ is increasing if and only if:

$$\frac{1}{(B+1)^2} > \left(\frac{2}{(B+1)^2} + \frac{10B-2+2\sqrt{9B^2-10B+1}}{\sqrt{9B^2-10B+1}(3B+\sqrt{9B^2-10B+1})^2}\right)\mu_0$$

$$\Leftrightarrow \quad \left[\sqrt{9B^2-10B+1}(3B+\sqrt{9B^2-10B+1})^2\right] \times \Bigg[2\sqrt{9B^2-10B+1}(3B+\sqrt{9B^2-10B+1})^2 + (B+1)^2(10B-2+2\sqrt{9B^2-10B+1})\Bigg]^{-1} > \mu_0$$

The left-hand side is clearly less than $\frac{1}{2}$, so there can be $\mu_0 \in \left(\frac{1}{B+1}, \frac{1}{2}\right)$ such that the inequality is not satisfied and $F_4(B,\mu_0)$ is decreasing in B. However, the left-hand side is greater than 0 and can be greater than $m_H = \frac{1}{B+1}$ for B large enough (e.g. at $B = 2$ the left-hand side is $\approx 0.39 > \frac{1}{3}$). Therefore, it is possible to find μ_0 and B such that $\bar{\mu} > \frac{1}{2}$ and $B > \underline{B}$ and such that $\frac{\partial F_4(B,\mu_0)}{\partial B} > 0$.

(b) **For $\boldsymbol{\mu_0}$:** Note that $F_4(B,\mu_0)$ can be re-written as a linear function of μ_0:

$$\begin{aligned} F_4(B,\mu_0) &= \frac{1}{(B+1)((3B-1)\mu_0+\mu_0\sqrt{9B^2-10B+1}+2)} \\ &\quad \times \Bigg[((B+3)\mu_0-B-2)\mu_0\sqrt{9B^2-10B+1} \\ &\qquad -2+(-B^2+6B-5)\mu_0^2+(3B^2-B+6)\mu_0\Bigg] \\ &= -\frac{1}{B+1} + \left(\frac{2B^2+8B+2\sqrt{9B^2-10B+1}-2}{(B+1)(3B+\sqrt{9B^2-10B+1}-1)}\right)\mu_0 \end{aligned}$$

It is easy to see that the slope is positive as $8B-2>0$ and $3B-1>0$:

$$\frac{2B^2+8B+2\sqrt{9B^2-10B+1}-2}{(B+1)(3B+\sqrt{9B^2-10B+1}-1)} > 0$$

So that $F_4(B,\mu_0)$ is increasing in μ_0 everywhere.

5. **Case 5:** If $\overline{B} < B$, $\frac{\partial F(B,\mu_0)}{\partial B} = -\frac{2\mu_0}{(B+1)^2} < 0$ and $\frac{\partial F(B,\mu_0)}{\partial \mu_0} = \frac{2}{B+1} > 0$.

6. Finally, the influence function has a discontinuity at $\mu_0 = m_H$ where influence drops,

so influence is non-monotonic in alignment. Indeed, when $B < \underline{B}$,

$$\lim_{\mu_0 \to m_H^-} F(B, m_H) > \lim_{\mu_0 \to m_H^+} F(B, m_H)$$

as

$$\lim_{\mu_0 \to m_H^-} F(B, m_H) = \frac{(B+1)m_H}{B} := F_1(B, m_H)$$

and

$$\lim_{\mu_0 \to m_H^+} F(B, m_H) = \frac{(3m_H - 1)B + m_H}{B(B+1)} := F_3(B, m_H)$$

and as

$$F_1(B, m_H) - F_3(B, m_H) = \frac{(B+1)m_H}{B} - \frac{(3m_H - 1)B + m_H}{B(B+1)} = \frac{2B}{B+1} > 0$$

Similarly, when $B > \underline{B}$, we can show that

$$\lim_{\mu_0 \to m_H^-} F(B, m_H) > \lim_{\mu_0 \to m_H^+} F(B, m_H)$$

Since

$$\begin{aligned} F_2(B, m_H) - F_4(B, m_H) &= \left(m_H + (1 - m_H) \frac{m^*(\bar{\mu})}{(1 - m^*(\bar{\mu}))} \right) \\ &\quad - \left(m_H + (1 - m_H) \frac{m^*(\bar{\mu})}{(1 - m^*(\bar{\mu}))} - \frac{m_H - \underline{\mu}}{\bar{\mu} - \underline{\mu}} \right) \\ &= \frac{2B}{(B+1)^2} > 0 \end{aligned}$$

For the statements on welfare, I simply note that on $\left[\underline{B}, \overline{B}\right]$, or on $\left[B_H, \overline{B}\right]$, the policy maker's utility is either $U^G(\underline{\mu}, \bar{\mu})$ or $U^G(m^*(\bar{\mu}), \bar{\mu})$. Both functions are increasing in B and decreasing in μ_0 (this follows from the proof of Proposition 4).

□

Proof of Proposition 21. 1. Consider first the case where $B < \overline{B}$, and consider the following equilibrium: the policy maker chooses p to induce beliefs $(\underline{\mu}, \bar{\mu})$ if $B < \underline{B}$ and $(m^*(\bar{\mu}), \bar{\mu})$ if $B > \underline{B}$, the policy maker reports $\hat{r} = \varnothing$ following both r_0 and r_1 and the lobbyist chooses π_G. The only relevant deviations to consider are for the policy maker to disclose $\hat{r} = r$ instead of $\hat{r} = \varnothing$. In that case, the lobbyist learns the policy maker's type and chooses π_G if $\hat{r} = r_0$ and π_T if $\hat{r} = r_1$. The expected utility of type r_0 is therefore the same whether she deviates or not so she has no incentives to deviate. The expected utility of type r_1 is lower if she deviates so she has strict incentives not to deviate.

2. Consider now the case where $B > \overline{B}$, and consider the following equilibrium: the policy maker chooses p to induce beliefs $(\underline{\mu}, \bar{\mu})$ and reports $\hat{r} = r$ following both r_0 and r_1 and the lobbyist chooses π_G following $\hat{r} = r_0$ and π_T following $\hat{r} = r_1$. Let $\rho = \mathbb{P}(r = r_1 | \hat{r} = \varnothing)$ the off-equilibrium beliefs of the lobbyist following no disclosure. If ρ is such that the incentive constraint is not satisfied: $\mathbb{P}_{\pi_G}(s_1) < \rho \mathbb{P}_{\pi_T}(s_1|r_1)$, then the lobbyist chooses π_T following no disclosure. This gives type r_1 the same utility as under full disclosure so she has no incentives to deviate, and gives type r_0 a lower utility than under full disclosure.

□

Proof of Proposition 22. To prove the statement, I provide a numerical example of an investigation with three signal realisations that increases the expected utility of the policy maker.

Consider first the optimal investigation when $\mu_0 = 0.2$ and $B = 3$. Since $\underline{B} = 1 + \sqrt{2}$, we have $B > \underline{B}$. Given $\mu_0 = 0.2$, we have $\overline{B} \simeq 13.50$, so $B < \overline{B}$. Therefore, by Proposition 3, the optimal investigation induces beliefs $\mu^{r_0} = m^*(\bar{\mu}) \simeq 0.0898$, and $\mu^{r_1} = \bar{\mu} \simeq 0.4286$. This yields an expected utility to the policy maker of:

$$U^G(m^*(\bar{\mu}), \bar{\mu}) = (1 - \mu_0)\frac{(1 - 2m^*(\bar{\mu}))}{(1 - m^*(\bar{\mu}))} + \mu_0 \simeq 0.9211$$

Consider now a preliminary investigation p with three realisations $\{r_0, r_*, r_1\}$ such that:

$$p(r_0|0) = 0.765$$

$$p(r_0|1) = 0.300$$

$$p(r_*|0) = 0.035$$

$$p(r_*|1) = 0.100$$

This investigation induces three possible posterior beliefs: $\underline{\mu} < \mu^{r_0} \simeq 0.0893 < m^*(\bar{\mu})$, $\mu^{r_1} = \bar{\mu}$, and $\underline{\mu} < \mu^{r_*} = 0.4167 < \bar{\mu}$.

The lobbyist prefers to target the sceptical type and get expected utility $\mathbb{P}_{\pi_{r_0}}(s_1) =$

0.278 than to target the sympathetic type and get $\mathbb{P}_{\pi_{r_1}}(s_1)\,\mathbb{P}(r_1|s_1) = 0.240 < 0.278$ or the middle type and get $\mathbb{P}_{\pi_{r_*}}(s_1)[\mathbb{P}(r_1|s_1) + \mathbb{P}(r_*|s_1) = 0.274 < 0.278$.

The policy maker's expected utility is $U_3^G(\mu^{r_0}, \mu^{r_*}, \mu^{r_1}) \simeq 0.9216 > U^G(m^*(\bar{\mu}), \bar{\mu}) \simeq 0.9211$, and she therefore gains from the investigation with 3 realisations.

□

B Proofs of results in Chapter 2

Proof of Proposition 7. On the equilibrium path, interim beliefs should belong to the set $\{\mu_{\theta_L}, \mu_0, \mu_{\theta_R}\}$ as only pure strategies are allowed. Off-equilibrium, the modified intuitive criterion refinement also rules out beliefs outside the set $\{\mu_{\theta_L}, \mu_0, \mu_{\theta_R}\}$ since either only one of two types would like to deviate, in which case the out-of-equilibrium beliefs following this deviation should put weight on this type only, or neither / both types want to deviate in which case the out-of-equilibrium should equal the prior.

Equilibrium of the competitive game In the proposed pooling equilibrium, the equilibrium interim belief of the receiver is μ_0, and the equilibrium experiment induces posterior beliefs $(\mu_s)_{(s\in\boldsymbol{\pi^c})}$. The senders expected utilities are $\mathbb{E}_{\mu_\theta}\left[V_I((\mu_s)_{(s\in\boldsymbol{\pi^c})};\mu_0)\right]$ for Sender I of type θ and $\mathbb{E}_{\mu_0}\left[V_N((\mu_s)_{(s\in\boldsymbol{\pi^c})};\mu_0)\right]$ for Sender N.

Conditions 2.2 requires the uninformed sender to prefer the equilibrium distribution τ^c to any deviation, given that a deviation by this sender cannot change the out-of-equilibrium beliefs of the receiver.

2.3 requires that both types of the informed sender prefer the equilibrium distribution τ^c to any deviation, given any out-of-equilibrium beliefs $\tilde{\mu}(\pi')$ that satisfy condition 2.4.

2.4 defines the set of beliefs that satisfy the intuitive criterion. The intuitive criterion requires that following a deviation π', the receiver only puts weight on types of the sender who would prefer that deviation to the equilibrium outcome, *given some* interim beliefs of the receiver (the deviation is not 'equilibrium-dominated').

In particular, if the out-of-equilibrium belief $\tilde{\mu}(\pi')$ corresponds to a type θ, then that type θ must get a higher payoff from π' than from π^c, given some interim beliefs of the receiver. This is ensured by the first part of condition 2.4:

$$\exists\, \mu \ \text{ s.t. } \mathbb{E}_{\mu_\theta}\left[V_I((\mu_s)_{(s\in\pi'\vee\pi^c)};\mu)\right] > \mathbb{E}_{\mu_\theta}\left[V_I((\mu_s)_{(s\in\boldsymbol{\pi^c})};\mu_0)\right]$$

In addition, since the modification of the intuitive criterion used here imposes that the receiver should have belief μ_0 whenever both types potentially find the deviation profitable (possibly under different interim beliefs of the receiver), then the receiver can have out-of-equilibrium belief μ_θ only if type θ' never finds that deviation profitable. This is ensured by the second part of condition 2.4:

$$\forall\ \mu,\ \ \mathbb{E}_{\mu_{\theta'}}\left[V_I((\mu_s)_{(s\in\pi'\vee\pi^c)};\mu)\right] < \mathbb{E}_{\mu_{\theta'}}\left[V_I((\mu_s)_{(s\in\boldsymbol{\pi^c})};\mu_0)\right]$$

Finally, if either both types would find that deviation profitable given some interim beliefs of the receiver, or if neither types do, then the intuitive criterion does not restrict the out-of-equilibrium beliefs of the receiver in any way. The modified intuitive criterion used here restricts the receiver to have beliefs equal to the prior: $\tilde{\mu}(\pi') = \mu_0$. This is the case when either of the last two parts of condition 2.4 are satisfied:

$$\begin{aligned}
&\forall\ \mu_\theta,\ \forall\ \mu,\ \ \mathbb{E}_{\mu_\theta}\left[V_I((\mu_s)_{(s\in\pi'\vee\pi^c)};\mu)\right] < \mathbb{E}_{\mu_\theta}\left[V_I((\mu_s)_{(s\in\boldsymbol{\pi^c})};\mu_0)\right]\\
\text{Or,}\quad &\exists\ \mu(\theta),\ \text{s.t.}\quad \mathbb{E}_{\mu_\theta}\left[V_I((\mu_s)_{(s\in\pi'\vee\pi^c)};\mu(\theta))\right] > \mathbb{E}_{\mu_\theta}\left[V_I((\mu_s)_{(s\in\boldsymbol{\pi^c})};\mu_0)\right]\ \forall\mu_\theta
\end{aligned}$$

Equilibrium of the collusive game In the proposed pooling equilibrium, the equilibrium interim belief of the receiver is μ_0, and the equilibrium experiment induces posterior beliefs $(\mu_s)_{(s\in\pi^m)}$. The merged senders' expected utility is $\mathbb{E}_{\mu_\theta}\left[V_m((\mu_s)_{(s\in\pi^m)};\mu_0)\right]$ when the coalition's type is θ.

Condition 2.5 ensures that the coalition would not deviate to an experiment π' if that deviation induces out-of-equilibrium beliefs $\tilde{\mu}(\pi')$.

Condition 2.6 defines the set of beliefs that satisfy the modified intuitive criterion following a deviation to experiment π' following the same logic as condition 2.4.

No less informative equilibria in collusion Conditions 2.7 ensures that at least one type of the coalition would deviate from any experiment that generates a less informative distribution of posterior beliefs, *for any* out-of-equilibrium belief of the receiver that satisfy the modified intuitive criterion. If this condition was not satisfied, then a less informative distribution of posterior beliefs could be sustained as an equilibrium of the collusive game. This could be achieved by imposing some out-of-equilibrium beliefs in the set of possible ones that make the deviation unattractive to that type of the sender.

Condition 2.8 defines the set of out-of-equilibrium beliefs that could be used to rule out such deviations.

$\square$

Proof of Lemma 3. I first define the experiments π^m and π^c leading to the desired distribution of posterior beliefs. Recall that an experiment π consists of a finite partition of $\{L, R\} \times [0, 1]$, such that $\pi \subset S$ where S is the set of non-empty Lebesgue measurable subsets of $\{L, R\} \times [0, 1]$, and that there is a random variable X independent of ω and uniformly distributed on $[0, 1]$, and an S-valued random variable equal to s when $(\omega, x) \in s$, for $\omega \in \{L, R\}$ and x, a realisation of X.

Therefore, as $\mathbb{P}(s|\omega) = \lambda(\{x|(\omega, x) \in s\})$ (where $\lambda(.)$ is the Lebesgue measure), we can construct π^m as:

$$\begin{aligned}\pi^m = \{s_1, s_2\} \text{ where } s_1 &= \left(L, \left[0, \frac{\mu_2 - \mu_0}{\mu_2(1-\mu_0)}\right]\right) \cup (R, \emptyset)\\ \text{and } s_2 &= \left(L, \left[\frac{\mu_2 - \mu_0}{\mu_2(1-\mu_0)}, 1\right]\right) \cup (R, [0, 1])\end{aligned}$$

Similarly, we can construct π^c as:

$$\begin{aligned}\pi^c = \{s_1, s_2\} \text{ where } s_1 &= \left(L, \left[0, \frac{(1-\mu_1)(\mu_2 - \mu_0)}{(1-\mu_0)(\mu_2-\mu_1)}\right]\right) \cup \left(R, \left[0, \frac{\mu_1(\mu_2 - \mu_0)}{\mu_0(\mu_2-\mu_1)}\right]\right)\\ \text{and } s_2 &= \left(L, \left[\frac{(1-\mu_1)(\mu_2 - \mu_0)}{(1-\mu_0)(\mu_2-\mu_1)}, 1\right]\right) \cup \left(R, \left[\frac{\mu_1(\mu_2 - \mu_0)}{\mu_0(\mu_2-\mu_1)}, 1\right]\right)\end{aligned}$$

As a result, $\pi^c \vee \pi^c$ has realisations $\{z_1, z_{12}, z_{21}, z_2\}$ where $z_1 = s_1 \cap s_1$, $z_{12} = s_1 \cap s_2$, $z_{21} = s_2 \cap s_1$, and $z_2 = s_1 \cap s_1$. Since $z_{12} = s_1 \cap s_2 = \emptyset$ and $z_{21} = s_2 \cap s_1 = \emptyset$, however, and since $z_1 = s_1 \cap s_1 = s_1$ and $z_2 = s_2 \cap s_2 = s_2$ the support of $\pi^c \vee \pi^c$ is $\{s_1, s_2\}$.

I now proceed to show why the two conditions are necessary:

1. If $\boldsymbol{\pi^c}$ is an equilibrium of the competitive game, then $w_i^1 < w_j^1$ for any i, j such that $i > j$.

   Suppose not, then for some $i, j \in \{1, 2, 3\}$, $w_i^1 > w_j^1$ and $i > j$. If $w_3^I > w_1^I$ or $w_3^I > w_2^I$, then type θ_R of Sender I would prefer to deviate to a fully disclosing experiment and get w_3^I instead of the combination of w_1^I and w_2^I obtained when both senders play π^c. Therefore, $w_3^I > w_2^I$, so given assumption 1, it must be that $w_2^I > w_1^I$.

2. If π^m is an equilibrium of the collusive game, then $\max\{w_3^c, w_1^c\} < w_2^c$.

   Suppose not, then $w_3^c > \tau_1^m w_1^c + \tau_2^m w_2^c$ and type θ_R of the coalition could deviate to a fully disclosing experiment and get w_3^c.

$\square$

Proof of Proposition 8. I first show that the conditions in 8 guarantee the existence of a competitive equilibrium inducing a distribution of posteriors τ^c, then that they guarantee the existence of a collusive equilibrium inducing a distribution of posteriors τ^m, and finally that there is no equilibrium in collusion inducing a less informative distribution than τ^c.

Equilibrium in competition

I show that the conditions in Proposition 8 are sufficient for each condition in Proposition 7 to be satisfied, given these possible deviations.

1. **Uninformed sender:**

   Claim 1: Condition 2.12 $\Rightarrow$ condition 1(a) in Proposition 8.

   Proof: Following Kamenica & Gentzkow (2011), we know that the optimal distribution of posterior beliefs should yield a payoff on the concave closure of the sender's utility. Given the shape of the sender's utility function, the concave closure below μ_2 is given by either the line joining points $(0,0)$ and (μ_1, w_1^N) or the line joining $(0,0)$ and (μ_2, w_2^N). Above μ_2, it is given by w_3^N.

   Thus, if $w_2^N < \frac{w_3^N}{\mu_2}\mu_1$, it is given by

$$\widetilde{V}(\mu) = \begin{cases} \frac{w_3^N}{\mu_2}\mu & \text{if } \mu \leq \mu_2 \\ w_3^N & \text{if } \mu > \mu_2 \end{cases}$$

   And if $w_2^N > \frac{w_3^N}{\mu_2}\mu_1$, it is given by

$$\widetilde{V}(\mu) = \begin{cases} \frac{w_2^N}{\mu_1}\mu & \text{if } \mu \leq \mu_1 \\ \left(\frac{\mu_2 w_2^N - \mu_1 w_3^N}{\mu_2 - \mu_1}\right) + \left(\frac{w_3^N - w_2^N}{\mu_2 - \mu_1}\right)\mu & \text{if } \mu \leq \mu_1 \\ w_3^N & \text{if } \mu > \mu_2 \end{cases}$$

   Therefore, condition 2.12: $\frac{\mu_1}{\mu_2} \leq \frac{w_2^N}{w_3^N}$ implies that the concave closure is given by the second expression, and given the prior $\mu_0 \in [\mu_1, \mu_2]$, the optimal experiment induces beliefs μ_1 and μ_2. That is,

$$\mathbb{E}_{\mu_0}[V_N(\mu_1, \mu_2; \mu_0)] \geq \mathbb{E}_{\mu_0}[V_N((\mu_s)_{(s \in \pi' \vee \pi^c)}; \mu_0)]$$

   Therefore, condition 1(a) of proposition 7 is satisfied given condition 2.12.

2. **Informed sender:**

Lemma 7. *Conditions 2.12 and 2.13 $\Rightarrow$ condition 1(b) in Proposition 8.*

Proof. **Claim 1:** If condition 2.13 is satisfied, then Sender I of type θ_L does not deviate from π^c if that deviation induces out-of-equilibrium beliefs μ_0.

Proof: I first show that condition 2.13 implies that type θ_L of Sender I prefers to induce (μ_1, μ_2) than $(\mu_1, 1)$.

$$\begin{aligned}
\mathbb{E}_{\mu_{\theta_L}}[V_I(\mu_1,\mu_2;\mu_0)] &= \left[(1-\mu_{\theta_L})\left(\frac{(1-\mu_1)(\mu_2-\mu_0)}{(1-\mu_0)(\mu_2-\mu_1)}\right) + \mu_{\theta_L}\left(\frac{\mu_1(\mu_2-\mu_0)}{\mu_0(\mu_2-\mu_1)}\right)\right] w_1^I \\
&\quad + \left[(1-\mu_{\theta_L})\left(1-\frac{(1-\mu_1)(\mu_2-\mu_0)}{(1-\mu_0)(\mu_2-\mu_1)}\right)\right. \\
&\quad \left. + \mu_{\theta_L}\left(1-\frac{\mu_1(\mu_2-\mu_0)}{\mu_0(\mu_2-\mu_1)}\right)\right] w_2^I \\
&\geq \left[(1-\mu_{\theta_L}) + \mu_{\theta_L}\frac{\mu_1(1-\mu_0)}{\mu_0(1-\mu_1)}\right] w_1^I \\
&\quad + \left(1-\left[(1-\mu_{\theta_L}) + \mu_{\theta_L}\frac{\mu_1(1-\mu_0)}{\mu_0(1-\mu_1)}\right]\right) 0 \\
&= \left[(1-\mu_{\theta_L}) + \mu_{\theta_L}\frac{\mu_1(1-\mu_0)}{\mu_0(1-\mu_1)}\right] w_1^I \\
&= \mathbb{E}_{\mu_{\theta_L}}[V_I(\mu_1,1;\mu_0)]
\end{aligned}$$

Re-arranging, we therefore, get

$$\begin{aligned}
& \mathbb{E}_{\mu_{\theta_L}}[V_I(\mu_1,\mu_2;\mu_0)] \geq \mathbb{E}_{\mu_{\theta_L}}[V_I(\mu_1,1;\mu_0)] \\
\Leftrightarrow \quad & \frac{(1-\mu_2)(\mu_0(1-\mu_1)-\mu_{\theta_L}(\mu_0-\mu_1))}{(1-\mu_1)(\mu_0(1-\mu_2)+\mu_{\theta_L}(\mu_2-\mu_0))} \leq \frac{w_2^I}{w_1^I}
\end{aligned}$$

Next, note that $\mathbb{E}_{\mu_{\theta_L}}[V_I(\mu_1,1;\mu_0)] \geq \mathbb{E}_{\mu_{\theta_L}}[V_I(m_1,m_2;\mu_0)]$ for any $m_1 \in [0,\mu_1]$ and $m_2 \in (\mu_2, 1]$ as inducing any belief $m_1 \in [0,\mu_1)$ or $m_2 \in (\mu_2,1]$ reduces the probability of receiving w_1^I.

So if $\mathbb{E}_{\mu_{\theta_L}}[V_I(\mu_1,\mu_2;\mu_0)] \geq \mathbb{E}_{\mu_{\theta_L}}[V_I(\mu_1,1;\mu_0)]$, then

$$\mathbb{E}_{\mu_{\theta_L}}[V_I(\mu_1,\mu_2;\mu_0)] \geq \mathbb{E}_{\mu_{\theta_L}}[V_I(m_1,m_2;\mu_0)]$$

for any $m_1 \in [0,\mu_1]$ and $m_2 \in [\mu_2, 1]$.

Similarly $\mathbb{E}_{\mu_{\theta_L}}[V_I(\mu_1,\mu_2;\mu_0)] \geq \mathbb{E}_{\mu_{\theta_L}}[V_I(m_1,m_2;\mu_0)]$ for any $m_1 \in [0,\mu_1]$ and $m_2 \in [\mu_0,\mu_2]$ as inducing these beliefs reduces the probability of receiving w_1^I.

Finally, $\mathbb{E}_{\mu_{\theta_L}}[V_I(\mu_1,\mu_2;\mu_0)] \geq \mathbb{E}_{\mu_{\theta_L}}[V_I(m_1,m_2;\mu_0)]$ for any $m_1 \in (\mu_1,\mu_0]$ and $m_2 \in [\mu_0,1]$ as this would only yields a mixture of w_1^N and 0 rather than a mixture of w_1^I and w_1^N, and since $w_1^I > w_1^N > 0$.

Claim 2: If $\frac{\mu_1}{1-\mu_1}\frac{1-\mu_2}{\mu_2} \leq \frac{w_2^I}{w_1^I}$, then Sender I of type θ_R does not deviate from π^c if that deviation induces out-of-equilibrium beliefs μ_0.

Proof: I first show that $\frac{\mu_1}{1-\mu_1}\frac{1-\mu_2}{\mu_2} \leq \frac{w_2^I}{w_1^I}$ implies that type θ_R of Sender I prefers to induce (μ_1,μ_2) than $(\mu_1,1)$.

$$\begin{aligned}
\mathbb{E}_{\mu_{\theta_R}}[V_I(\mu_1,\mu_2;\mu_0)] &= \left[(1-\mu_{\theta_R})\left(\frac{(1-\mu_1)(\mu_2-\mu_0)}{(1-\mu_0)(\mu_2-\mu_1)}\right) + \mu_{\theta_R}\left(\frac{\mu_1(\mu_2-\mu_0)}{\mu_0(\mu_2-\mu_1)}\right)\right] w_1^I \\
&\quad + \Bigg[(1-\mu_{\theta_R})\left(1-\frac{(1-\mu_1)(\mu_2-\mu_0)}{(1-\mu_0)(\mu_2-\mu_1)}\right) \\
&\quad + \mu_{\theta_R}\left(1-\frac{\mu_1(\mu_2-\mu_0)}{\mu_0(\mu_2-\mu_1)}\right)\Bigg] w_2^I \\
&= \left(\frac{\mu_1(\mu_2-\mu_0)}{\mu_0(\mu_2-\mu_1)}\right) w_1^I + \left(1-\frac{\mu_1(\mu_2-\mu_0)}{\mu_0(\mu_2-\mu_1)}\right) w_2^I \\
&\geq \frac{\mu_1(1-\mu_0)}{\mu_0(1-\mu_1)} w_1^I + \left(1-\frac{\mu_1(1-\mu_0)}{\mu_0(1-\mu_1)}\right) 0 \\
&= \frac{\mu_1(1-\mu_0)}{\mu_0(1-\mu_1)} w_1^I \\
&= \mathbb{E}_{\mu_{\theta_R}}[V_I(\mu_1,1;\mu_0)]
\end{aligned}$$

Re-arranging, we therefore, get

$$\mathbb{E}_{\mu_{\theta_R}}[V_I(\mu_1,\mu_2;\mu_0)] \geq \mathbb{E}_{\mu_{\theta_R}}[V_I(\mu_1,1;\mu_0)] \Leftrightarrow \frac{\mu_1}{1-\mu_1}\frac{1-\mu_2}{\mu_2} \leq \frac{w_2^I}{w_1^I}$$

Next, note that $\mathbb{E}_{\mu_{\theta_R}}[V_I(\mu_1,1;\mu_0)] \geq \mathbb{E}_{\mu_{\theta_R}}[V_I(m_1,m_2;\mu_0)]$ for any $m_1 \in [0,\mu_1]$ and $m_2 \in (\mu_2,1]$ as inducing any belief $m_1 \in [0,\mu_1)$ or $m_2 \in (\mu_2,1]$ reduces the probability of receiving w_1^I.

So if $\mathbb{E}_{\mu_{\theta_R}}[V_I(\mu_1,\mu_2;\mu_0)] \geq \mathbb{E}_{\mu_{\theta_R}}[V_I(\mu_1,1;\mu_0)]$, then

$$\mathbb{E}_{\mu_{\theta_R}}[V_I(\mu_1,\mu_2;\mu_0)] \geq \mathbb{E}_{\mu_{\theta_R}}[V_I(m_1,m_2;\mu_0)]$$

for any $m_1 \in [0,\mu_1]$ and $m_2 \in [\mu_2,1]$.

Similarly $\mathbb{E}_{\mu_{\theta_R}}[V_I(\mu_1,\mu_2;\mu_0)] \geq \mathbb{E}_{\mu_{\theta_R}}[V_I(m_1,m_2;\mu_0)]$ for any $m_1 \in [0,\mu_1]$ and $m_2 \in [\mu_0,\mu_2]$ as inducing these beliefs reduces the probability of receiving w_1^I.

Finally, $\mathbb{E}_{\mu_{\theta_R}}[V_I(\mu_1,\mu_2;\mu_0)] \geq \mathbb{E}_{\mu_{\theta_R}}[V_I(m_1,m_2;\mu_0)]$ for any $m_1 \in (\mu_1,\mu_0]$ and $m_2 \in$

$[\mu_0, 1]$ as this would only yields a mixture of w_1^N and 0 rather than a mixture of w_1^I and w_1^N, and since $w_1^I > w_1^N > 0$.

Claim 3: Condition 2.13 $\Rightarrow$ Sender I of type θ_R does not deviate from π^c if that deviation induces out-of-equilibrium beliefs μ_0.

Proof: I show that condition 2.13 $\Rightarrow \frac{\mu_1}{1-\mu_1}\frac{1-\mu_2}{\mu_2} \leq \frac{w_2^I}{w_1^I}$.

$$\frac{(1-\mu_2)(\mu_0(1-\mu_1)-\mu_{\theta_L}(\mu_0-\mu_1))}{(1-\mu_1)(\mu_0(1-\mu_2)+\mu_{\theta_L}(\mu_2-\mu_0))} \geq \frac{\mu_1}{1-\mu_1}\frac{1-\mu_2}{\mu_2}$$
$$\Leftrightarrow \quad \mu_2(\mu_0(1-\mu_1)-\mu_{\theta_L}(\mu_0-\mu_1)) \geq \mu_1(\mu_0(1-\mu_2)+\mu_{\theta_L}(\mu_2-\mu_0))$$
$$\Leftrightarrow \quad \mu_2 \geq \mu_1$$

Therefore, condition 2.13 $\Rightarrow \frac{w_2^I}{w_1^I} \geq \frac{(1-\mu_2)(\mu_0(1-\mu_1)-\mu_{\theta_L}(\mu_0-\mu_1))}{(1-\mu_1)(\mu_0(1-\mu_2)+\mu_{\theta_L}(\mu_2-\mu_0))} \geq \frac{\mu_1}{1-\mu_1}\frac{1-\mu_2}{\mu_2}$, and using claim 2, Sender I of type θ_R does not deviate from π^c if that deviation induces out-of-equilibrium beliefs μ_0.

Claim 4: No sender wants to deviate to an experiment if that experiment induces out-of-equilibrium beliefs μ_{θ_R}.

Proof: When the receiver's out-of-equilibrium belief following some deviation is μ_{θ_R}, then the receiver's only possible posterior belief is $\mu = 1$. Therefore, Sender I's expected utility is w_3^I and since $w_3^I < w_2^I < w_1^I$, then any deviation is equilibrium dominated for both types of Sender I if it induces out-of-equilibrium beliefs μ_{θ_R}.

Claim 5: Condition 2.13 $\Rightarrow$ if a feasible deviation were to induce out-of-equilibrium belief μ_{θ_L}, type θ_L would not want to deviate to it.

Proof: I begin to prove this claim with a lemma defining the set of beliefs that are feasible, given that a deviation induces belief μ_{θ_L} and given that Sender N is playing π^c. I then prove a second lemma that shows that among any feasible deviation is weakly dominated by either one of two feasible deviations which I define. Finally, I show that Condition 2.13 ensures that type θ_L deviates to neither of these two deviations.

Lemma 8. *If the receiver's belief following a deviation from (π^c, π^c) is θ_L, then it must induce some belief $\mu \in [0, m(\mu_1, \mu_0, \mu_{\theta_L})] \cup [m(\mu_2, \mu_0, \mu_{\theta_L}), 1]$ with some positive probability.*

Proof of Lemma 8. By definition of the function $m(\mu, \mu_r, \mu_i)$, the pair of beliefs induced by π^c when the receiver's interim belief is μ_{θ_L} is $(m(\mu_1, \mu_0, \mu_{\theta_L}), m(\mu_2, \mu_0, \mu_{\theta_L}))$.

Consider a deviation from Sender N to π'. Suppose, by contradiction, that all beliefs induced by the joint experiment $\pi' \vee \pi^c$ are in the interval

$$(m(\mu_1, \mu_0, \mu_{\theta_L}), m(\mu_2, \mu_0, \mu_{\theta_L}))$$

Then $\langle \pi' \vee \pi^c \rangle$ is more integral-precise than $\langle \pi^c \rangle$ (intuitively, a distribution of posterior beliefs is more integral-precise if the posteriors in its support are less spread-out, see Ganuza & Penalva (2010) for the full definition). In addition, Ganuza & Penalva (2010) (Theorem 2) show that a distribution of posterior is more integral-precise than another if and only if it is Blackwell less informative. Therefore, if all beliefs induced by $\pi' \vee \pi^c$ are in $(m(\mu_1, \mu_0, \mu_{\theta_L}), m(\mu_2, \mu_0, \mu_{\theta_L}))$, then $\langle \pi^c \rangle$ must be strictly Blackwell more informative than $\langle \pi' \vee \pi^c \rangle$.

However, this contradicts the result that $\langle \pi' \vee \pi^c \rangle \succsim \langle \pi^c \rangle$ as $\pi' \vee \pi^c$ is a refinement of π^c (Gentzkow & Kamenica (2017*a*), Lemma 3). Therefore, any deviation to an experiment π' must induce some belief $\mu \in [0, m(\mu_1, \mu_0, \mu_{\theta_L})] \cup [m(\mu_2, \mu_0, \mu_{\theta_L}), 1]$ with some positive probability. $\square$

Lemma 9. *The most profitable feasible deviation given that Sender N plays π^c and given out-of-equilibrium belief θ_L induces either*

$$(m(\mu_1, \mu_0, \mu_{\theta_L}), \mu_2) \quad \textit{or} \quad (m(\mu_1, \mu_0, \mu_{\theta_L}), 1)$$

Proof of Lemma 9. This result follows a standard concavification argument. First, note that any distribution over posterior induced by some deviation that puts weight on more than 3 distinct beliefs can be replicated by a payoff-equivalent distribution that puts weight on at most 3 distinct beliefs. To do this, we can take any subset of beliefs in the support that results in the same payoff w and replace it by a single belief equal to the average of these beliefs. That is, for any $\mu' \in \text{supp}(\langle \pi' \vee \pi^c \rangle)$ such that $v_I(\mu') = w_i^1$, replace μ' by $\tilde{\mu} = \sum_{\{\mu \text{ s.t. } v_I(\mu) = w_i^1\}} \mathbb{P}(\mu)\mu$.

If any deviation induces beliefs that are not on the concave closure of $V_I(\mu)$ restricted to the set of feasible beliefs, then there is an alternative deviation that yields a higher payoff. Therefore, the best deviation induces at most two beliefs.

Finally, among the distribution of beliefs inducing two posterior beliefs, the best possible deviation must induce the feasible belief closest to the prior when the receiver updates downwards, that is $\mu = m(\mu_1, \mu_0, \mu_{\theta_L})$, and the belief furthest from the prior inducing either a payoff of w_2^I or 0, depending on the value of w_1^I and w_2^I. That is, either $\mu = \mu_2$ or $\mu = 1$.

Since both $\mu_2 > m(\mu_2, \mu_0, \mu_{\theta_L})$ and $1 > m(\mu_2, \mu_0, \mu_{\theta_L})$ (as $\mu_{\theta_L} < \mu_0$), then both deviations are feasible. $\square$

The next step is to show that condition 2.13 implies

$$\mathbb{E}_{\mu_{\theta_L}}[V_I(\mu_1, \mu_2; \mu_0)] \geq \mathbb{E}_{\mu_{\theta_L}}[V_I(m(\mu_1, \mu_0, \mu_{\theta_L}), \mu_2; \mu_{\theta_L})]$$

This holds since:

$$\begin{aligned}
\mathbb{E}_{\mu_{\theta_L}}[V_I(\mu_1, \mu_2; \mu_0)] = &\left[(1-\mu_{\theta_L})\left(\frac{(1-\mu_1)(\mu_2-\mu_0)}{(1-\mu_0)(\mu_2-\mu_1)}\right) + \mu_{\theta_L}\left(\frac{\mu_1(\mu_2-\mu_0)}{\mu_0(\mu_2-\mu_1)}\right)\right] w_1^I \\
&+ \left[(1-\mu_{\theta_L})\left(1 - \frac{(1-\mu_1)(\mu_2-\mu_0)}{(1-\mu_0)(\mu_2-\mu_1)}\right)\right. \\
&\left. + \mu_{\theta_L}\left(1 - \frac{\mu_1(\mu_2-\mu_0)}{\mu_0(\mu_2-\mu_1)}\right)\right] w_2^I \\
\geq &\left[(1-\mu_{\theta_L})\left(\frac{(1-m(\mu_1,\mu_0,\mu_{\theta_L}))(\mu_2-\mu_{\theta_L})}{(1-\mu_{\theta_L})(\mu_2-m(\mu_1,\mu_0,\mu_{\theta_L}))}\right)\right. \\
&\left. + \mu_{\theta_L}\left(\frac{m(\mu_1,\mu_0,\mu_{\theta_L})(\mu_2-\mu_{\theta_L})}{\mu_{\theta_L})(\mu_2-m(\mu_1,\mu_0,\mu_{\theta_L}))}\right)\right] \\
&\times \left(\frac{m(\mu_1,\mu_0,\mu_{\theta_L})}{\mu_1}\right) w_1^I \\
&+ \left(1 - \left[(1-\mu_{\theta_L})\left(\frac{(1-m(\mu_1,\mu_0,\mu_{\theta_L}))(\mu_2-\mu_{\theta_L})}{(1-\mu_{\theta_L})(\mu_2-m(\mu_1,\mu_0,\mu_{\theta_L}))}\right)\right.\right. \\
&\left.\left. + \mu_{\theta_L}\left(\frac{m(\mu_1,\mu_0,\mu_{\theta_L})(\mu_2-\mu_{\theta_L})}{\mu_{\theta_L}(\mu_2-m(\mu_1,\mu_0,\mu_{\theta_L}))}\right)\right]\right) w_2^I \\
= &\ \mathbb{E}_{\mu_{\theta_L}}[V_I(m(\mu_1,\mu_0,\mu_{\theta_L}), \mu_2; \mu_{\theta_L})]
\end{aligned}$$

This holds if and only if

$$
\begin{aligned}
&w_1^I\Bigg[\frac{(\mu_2-\mu_0)\left(\mu_0(1-\mu_{\theta_L})(1-\mu_1)+(1-\mu_0)\mu_{\theta_L}\mu_1\right)}{\mu_0(1-\mu_0)(\mu_2-\mu_1)} \\
&\qquad -\left(\frac{\mu_2-\mu_{\theta_L}}{\mu_2-m(\mu_1,\mu_0,\mu_{\theta_L})}\right)\left(\frac{m(\mu_1,\mu_0,\mu_{\theta_L})}{\mu_1}\right)\Bigg] \\
&\geq w_2^I\left[\frac{(\mu_2-\mu_0)\left(\mu_0(1-\mu_{\theta_L})(1-\mu_1)+(1-\mu_0)\mu_{\theta_L}\mu_1\right)}{\mu_0(1-\mu_0)(\mu_2-\mu_1)}-\left(\frac{\mu_2-\mu_{\theta_L}}{\mu_2-m(\mu_1,\mu_0,\mu_{\theta_L})}\right)\right]
\end{aligned}
$$

Re-arranging and substituting for $m(\mu_1,\mu_0,\mu_{\theta_L})$ shows that this is equivalent to:

$$
\begin{aligned}
&\frac{G_1(\mu_0,\mu_1,\mu_2,\mu_{\theta_L})\mu_1(\mu_2-m(\mu_1,\mu_0,\mu_{\theta_L}))-m(\mu_1,\mu_0,\mu_{\theta_L})(\mu_2-\mu_{\theta_L})(\mu_2-\mu_1)\mu_0(1-\mu_0)}{G_1(\mu_0,\mu_1,\mu_2,\mu_{\theta_L})\mu_1(\mu_2-m(\mu_1,\mu_0,\mu_{\theta_L}))-\mu_1(\mu_2-\mu_{\theta_L})(\mu_2-\mu_1)\mu_0(1-\mu_0)} \\
&\leq \frac{w_2^I}{w_1^I}
\end{aligned}
$$

Where

$$
G_1(\mu_0,\mu_1,\mu_2,\mu_{\theta_L})=(\mu_2-\mu_0)(\mu_0(1-\mu_1)(1-\mu_{\theta_L})+(1-\mu_0)\mu_{\theta_L}\mu_1)
$$

Finally, it is easy to verify that

$$
\begin{aligned}
&\frac{G_1(\mu_0,\mu_1,\mu_2,\mu_{\theta_L})\mu_1(\mu_2-m(\mu_1,\mu_0,\mu_{\theta_L}))-m(\mu_1,\mu_0,\mu_{\theta_L})(\mu_2-\mu_{\theta_L})(\mu_2-\mu_1)\mu_0(1-\mu_0)}{G_1(\mu_0,\mu_1,\mu_2,\mu_{\theta_L})\mu_1(\mu_2-m(\mu_1,\mu_0,\mu_{\theta_L}))-\mu_1(\mu_2-\mu_{\theta_L})(\mu_2-\mu_1)\mu_0(1-\mu_0)} \\
&\leq \frac{(1-\mu_2)(\mu_0(1-\mu_1)-\mu_{\theta_L}(\mu_0-\mu_1))}{(1-\mu_1)(\mu_0(1-\mu_2)+\mu_{\theta_L}(\mu_2-\mu_0))}
\end{aligned}
$$

Where the right-hand side of this inequality is the left-hand side of condition 2.13. So if 2.13 is satisfied, then this condition is satisfied.

Finally, I show that condition 2.13 also implies that

$$
\mathbb{E}_{\mu_{\theta_L}}[V_I(\mu_1,\mu_2;\mu_0)]\geq\mathbb{E}_{\mu_{\theta_L}}[V_I(m(\mu_1,\mu_0,\mu_{\theta_L}),1;\mu_{\theta_L})]
$$

This holds since:

$$\begin{aligned}\mathbb{E}_{\mu_{\theta_L}}[V_I(\mu_1,\mu_2;\mu_0)] &= \left[(1-\mu_{\theta_L})\left(\frac{(1-\mu_1)(\mu_2-\mu_0)}{(1-\mu_0)(\mu_2-\mu_1)}\right)+\mu_{\theta_L}\left(\frac{\mu_1(\mu_2-\mu_0)}{\mu_0(\mu_2-\mu_1)}\right)\right]w_1^I \\ &\quad + \left[(1-\mu_{\theta_L})\left(1-\frac{(1-\mu_1)(\mu_2-\mu_0)}{(1-\mu_0)(\mu_2-\mu_1)}\right)\right. \\ &\quad \left. + \mu_{\theta_L}\left(1-\frac{\mu_1(\mu_2-\mu_0)}{\mu_0(\mu_2-\mu_1)}\right)\right]w_2^I \\ &\geq \left[(1-\mu_{\theta_L})+\mu_{\theta_L}\frac{m(\mu_1,\mu_0,\mu_{\theta_L})(1-\mu_{\theta_L})}{\mu_{\theta_L}(1-m(\mu_1,\mu_0,\mu_{\theta_L}))}\right] \\ &\quad \times\left(\frac{m(\mu_1,\mu_0,\mu_{\theta_L})}{\mu_1}\right)w_1^I \\ &\quad + \left(1-\left[(1-\mu_{\theta_L})+\mu_{\theta_L}\frac{m(\mu_1,\mu_0,\mu_{\theta_L})(1-\mu_{\theta_L})}{\mu_{\theta_L}(1-m(\mu_1,\mu_0,\mu_{\theta_L}))}\right]\right)0 \\ &= \mathbb{E}_{\mu_{\theta_L}}[V_I(m(\mu_1,\mu_0,\mu_{\theta_L}),1;\mu_{\theta_L})]\end{aligned}$$

Note that, since $\frac{m(\mu_1,\mu_0,\mu_{\theta_L})}{\mu_1}<1$, this holds whenever the following holds:

$$\begin{aligned}&\left[(1-\mu_{\theta_L})\left(\frac{(1-\mu_1)(\mu_2-\mu_0)}{(1-\mu_0)(\mu_2-\mu_1)}\right)+\mu_{\theta_L}\left(\frac{\mu_1(\mu_2-\mu_0)}{\mu_0(\mu_2-\mu_1)}\right)\right]w_1^I \\ &\quad + \left[(1-\mu_{\theta_L})\left(1-\frac{(1-\mu_1)(\mu_2-\mu_0)}{(1-\mu_0)(\mu_2-\mu_1)}\right)+\mu_{\theta_L}\left(1-\frac{\mu_1(\mu_2-\mu_0)}{\mu_0(\mu_2-\mu_1)}\right)\right]w_2^I \\ &\geq \left[(1-\mu_{\theta_L})+\mu_{\theta_L}\frac{m(\mu_1,\mu_0,\mu_{\theta_L})(1-\mu_{\theta_L})}{\mu_{\theta_L}(1-m(\mu_1,\mu_0,\mu_{\theta_L}))}\right]w_1^I \\ &\quad + \left(1-\left[(1-\mu_{\theta_L})+\mu_{\theta_L}\frac{m(\mu_1,\mu_0,\mu_{\theta_L})(1-\mu_{\theta_L})}{\mu_{\theta_L}(1-m(\mu_1,\mu_0,\mu_{\theta_L}))}\right]\right)0\end{aligned}$$

Which holds if and only if

$$\begin{aligned}&w_2^I\left[\frac{(\mu_0-\mu_1)(\mu_{\theta_L}(\mu_2-\mu_0)+\mu_0(1-\mu_2))}{\mu_0(1-\mu_0)(\mu_2-\mu_1)}\right] \\ &\quad \geq w_1^I\left[\frac{(\mu_0-\mu_1)(\mu_{\theta_L}(\mu_2-\mu_0)+\mu_0(1-\mu_2))}{\mu_0(1-\mu_0)(\mu_2-\mu_1)}-\left(\frac{1-\mu_{\theta_L}}{1-m(\mu_1,\mu_0,\mu_{\theta_L})}\right)\right]\end{aligned}$$

Re-arranging and substituting for $m(\mu_1,\mu_0,\mu_{\theta_L})$ shows that this is equivalent to:

$$\frac{w_2^I}{w_1^I}\geq\frac{(1-\mu_2)(\mu_0(1-\mu_1)-\mu_{\theta_L}(\mu_0-\mu_1))}{(1-\mu_1)(\mu_0(1-\mu_2)+\mu_{\theta_L}(\mu_2-\mu_0))}$$

Which is condition 2.13.

□

Equilibrium in collusion

Since any deviations induces out-of-equilibrium belief μ_{θ_L}, the set of feasible beliefs that can be induced by these deviations is $[0, \mu_{\theta_L}] \cup [\mu_{\theta_L}, 1]$.

Following a standard concavification argument, the best possible deviations given this set induce either (m_1, μ_1) or (m_1, μ_2) for some $m_1 \in [0, \mu_{\theta_L}]$.

I show that if condition 2.14 is satisfied, then type θ_L of the coalition does not want to deviate to induce any of these pairs of beliefs.

If condition 2.14 is satisfied, then $\mathbb{E}_{\mu_{\theta_L}}[V_m(0, \mu_2; \mu_0)] \geq \mathbb{E}_{\mu_{\theta_L}}[V_m(m_1, \mu_2; \mu_{\theta_L})]$ for any $m_1 \in [0, \mu_{\theta_L}]$. This holds since:

$$\begin{aligned}
\mathbb{E}_{\mu_{\theta_L}}[V_m(0, \mu_2; \mu_0)] &= \left(\frac{(1-\mu_{\theta_L})(\mu_2 - \mu_0)}{\mu_2(1-\mu_0)}\right) 0 \\
&\quad + \left(\frac{\mu_{\theta_L}\mu_2(1-\mu_0) + \mu_0(1-\mu_2)(1-\mu_{\theta_L})}{\mu_2(1-\mu_0)}\right) w^c(\mu_2) \\
&\geq \left(\frac{\mu_2 - \mu_{\theta_L}}{\mu_2 - m_1}\right)\left(\frac{w_1^I}{\mu_1} m_1\right) + \left(\frac{\mu_{\theta_L} - m_1}{\mu_2 - m_1}\right) w^c(\mu_2) \\
&= \mathbb{E}_{\mu_{\theta_L}}[V_m(m_1, \mu_2; \mu_{\theta_L})]
\end{aligned}$$

Which holds if and only if

$$\begin{aligned}
&\left[\frac{(\mu_{\theta_L}\mu_2(1-\mu_0) + \mu_0(1-\mu_2)(1-\mu_{\theta_L}))(\mu_2 - m_1) - (\mu_{\theta_L} - m_1)\mu_2(1-\mu_0)}{\mu_2(1-\mu_0)}\right] w^c(\mu_2) \\
&\geq \left(\frac{m_1(\mu_2 - \mu_{\theta_L})}{\mu_1}\right) w_1^I
\end{aligned}$$

Note that

$$\begin{aligned}
&(\mu_{\theta_L}\mu_2(1-\mu_0) + \mu_0(1-\mu_2)(1-\mu_{\theta_L}))(\mu_2 - m_1) - (\mu_{\theta_L} - m_1)\mu_2(1-\mu_0) \\
&= \mu_2\left(\mu_{\theta_L}\mu_2(1-\mu_0) + \mu_0(1-\mu_2)(1-\mu_{\theta_L}) - \mu_{\theta_L}(1-\mu_0)\right) \\
&\qquad + m_1\left[\mu_2(1-\mu_0) - (\mu_{\theta_L}\mu_2(1-\mu_0) + \mu_0(1-\mu_2)(1-\mu_{\theta_L}))\right] \\
&= \mu_2(1-\mu_2)(\mu_0 - \mu_{\theta_L}) + m_1(1-\mu_{\theta_L})(\mu_2 - \mu_0)
\end{aligned}$$

Which is positive for any $m_1 \in [0, \mu_{\theta_L}]$. Therefore, the condition for no deviation becomes

$$\frac{w^c(\mu_2)}{w_1^N} \geq \frac{m_1\mu_2(1-\mu_0)(\mu_2 - \mu_{\theta_L})}{\mu_1\left[\mu_2(1-\mu_2)(\mu_0 - \mu_{\theta_L}) + m_1(1-\mu_{\theta_L})(\mu_2 - \mu_0)\right]}$$

The right-hand side is increasing in m_1, so for this to be satisfied for any $m_1 \in [0, \mu_{\theta_L}]$, we need it to be satisfied at $m_1 = \mu_{\theta_L}$, that is:

$$\begin{aligned}
\frac{w^c(\mu_2)}{w_1^N} &\geq \frac{\mu_{\theta_L}\mu_2(1-\mu_0)(\mu_2-\mu_{\theta_L})}{\mu_1\left[\mu_2(1-\mu_2)(\mu_0-\mu_{\theta_L})+\mu_{\theta_L}(1-\mu_{\theta_L})(\mu_2-\mu_0)\right]} \\
&= \frac{\mu_{\theta_L}\mu_2(1-\mu_0)(\mu_2-\mu_{\theta_L})}{\mu_1(\mu_2-\mu_{\theta_L})\left[\mu_2\mu_{\theta_L}(1-\mu_0)+\mu_0(1-\mu_{\theta_L})(1-\mu_2)\right]} \\
&= \frac{\mu_{\theta_L}\mu_2(1-\mu_0)}{\mu_1\left[\mu_2\mu_{\theta_L}(1-\mu_0)+\mu_0(1-\mu_{\theta_L})(1-\mu_2)\right]}
\end{aligned}$$

Finally, if 2.14 is satisfied, then since $w^c(\mu_1) > w_1^I$,

$$\begin{aligned}
\frac{w^c(\mu_2)}{w_1^I} &\geq \frac{w^c(\mu_2)}{w^c(\mu_1)} \\
&\geq \frac{\mu_{\theta_L}\mu_2(1-\mu_0)}{\mu_1(\mu_{\theta_L}\mu_2(1-\mu_0)+\mu_0(1-\mu_{\theta_L})(1-\mu_2))}
\end{aligned}$$

I now show that if condition 2.14 is satisfied, then

$$\mathbb{E}_{\mu_{\theta_L}}[V_m(0,\mu_2;\mu_0)] \geq \mathbb{E}_{\mu_{\theta_L}}[V_m(m_1,\mu_1;\mu_{\theta_L})]$$

for any $m_1 \in [0, \mu_{\theta_L}]$. This holds since:

$$\begin{aligned}
\mathbb{E}_{\mu_{\theta_L}}[V_m(0,\mu_2;\mu_0)] &= \left(\frac{(1-\mu_{\theta_L})(\mu_2-\mu_0)}{\mu_2(1-\mu_0)}\right)0 \\
&\quad + \left(\frac{\mu_{\theta_L}\mu_2(1-\mu_0)+\mu_0(1-\mu_2)(1-\mu_{\theta_L})}{\mu_2(1-\mu_0)}\right)w^c(\mu_2) \\
&\geq \left(\frac{\mu_1-\mu_{\theta_L}}{\mu_1-m_1}\right)\left(\frac{w_1^I}{\mu_1}m_1\right)+\left(\frac{\mu_{\theta_L}-m_1}{\mu_1-m_1}\right)w^c(\mu_1) \\
&= \mathbb{E}_{\mu_{\theta_L}}[V_m(m_1,\mu_1;\mu_{\theta_L})]
\end{aligned}$$

This holds if and only if

$$\begin{aligned}
&\left(\frac{\mu_{\theta_L}\mu_2(1-\mu_0)+\mu_0(1-\mu_2)(1-\mu_{\theta_L})}{\mu_2(1-\mu_0)}\right)w^c(\mu_2) \\
&\quad \geq \left(\frac{\mu_1-\mu_{\theta_L}}{\mu_1-m_1}\right)\left(\frac{m_1}{\mu_1}w_1^I\right)+\left(\frac{\mu_{\theta_L}-m_1}{\mu_1-m_1}\right)w^c(\mu_1)
\end{aligned}$$

Note that $\left(\frac{\mu_1-\mu_{\theta_L}}{\mu_1-m_1}\right)\left(\frac{m_1}{\mu_1}\right)$ is increasing in m_1 and that $\frac{\mu_{\theta_L}-m_1}{\mu_1-m_1}$ is decreasing in m_1. In addition, recall that $w^c(\mu_1) > w_1^I$. Therefore, the right-hand side of the inequality is decreasing in m_1. So for it to be true for any $m_1 \in [0, \mu_{\theta_L}]$, we need to be true at $m_1 = 0$, that is

$$\left(\frac{\mu_{\theta_L}\mu_2(1-\mu_0)+\mu_0(1-\mu_2)(1-\mu_{\theta_L})}{\mu_2(1-\mu_0)}\right) w^c(\mu_2) \geq \left(\frac{\mu_{\theta_L}}{\mu_1}\right) w^c(\mu_1)$$

Which is equivalent to condition 2.14.

No less informative equilibrium in collusion

I show that, when condition 2.15 is satisfied, there is no equilibrium in collusion that induces a less informative distribution of beliefs than the competitive equilibrium τ^c. I proceed in three steps.

Claim 1: Condition 2.15 ensures that type θ_R would always prefer to deviate to a fully revealing experiment if both types of the coalition were to pool on an experiment inducing beliefs (μ_1, μ_2). As a result, a distribution over these two posteriors cannot be an equilibrium outcome of the collusive game.

Proof: Type θ_R obtains expected utility w_3^c with probability 1 with a fully revealing experiment. With an experiment inducing beliefs (μ_1, μ_2) given interim receiver belief μ_0, type θ_R gets expected utility:

$$\begin{aligned}
\mathbb{E}_{\mu_{\theta_R}}[V_c(\mu_1,\mu_2;\mu_0)] = & \left[(1-\mu_{\theta_R})\left(\frac{(1-\mu_1)(\mu_2-\mu_0)}{(1-\mu_0)(\mu_2-\mu_1)}\right) + \mu_{\theta_R}\left(\frac{\mu_1(\mu_2-\mu_0)}{\mu_0(\mu_2-\mu_1)}\right)\right] w^c(\mu_1) \\
& + \Bigg[(1-\mu_{\theta_R})\left(1-\frac{(1-\mu_1)(\mu_2-\mu_0)}{(1-\mu_0)(\mu_2-\mu_1)}\right) \\
& + \mu_{\theta_R}\left(1-\frac{\mu_1(\mu_2-\mu_0)}{\mu_0(\mu_2-\mu_1)}\right)\Bigg] w^c(\mu_2) \\
= & \left(\frac{\mu_1(\mu_2-\mu_0)}{\mu_0(\mu_2-\mu_1)}\right) w^c(\mu_1) + \left(1-\frac{\mu_1(\mu_2-\mu_0)}{\mu_0(\mu_2-\mu_1)}\right) w^c(\mu_2) \\
< & \; w_3^c = \mathbb{E}_{\mu_{\theta_R}}[V_I(0,1;\mu)]
\end{aligned}$$

Re-arranging, we therefore, get

$$\mathbb{E}_{\mu_{\theta_R}}[V_c(\mu_1,\mu_2;\mu_0)] < \mathbb{E}_{\mu_{\theta_R}}[V_c(0,1;\mu)] \Leftrightarrow \frac{\mu_1(\mu_2-\mu_0)}{\mu_0(\mu_2-\mu_1)} > \frac{w^c(\mu_2)-w_3^c}{w^c(\mu_2)-w^c(\mu_1)}$$

Which is condition 2.15.

Claim 2: There cannot be any less informative pooling equilibrium in collusion.

Proof: Any such equilibrium would need to induce some beliefs $m \in [\mu_1, \mu_2]$ (see Lemma 8). But in this case, the coalition could deviate to an experiment π' that induces beliefs $(0, \mu_1)$ given out-of-equilibrium belief μ_0, instead of m and keeps the distribution over other beliefs the same.

Note that since the deviation π' needs to induce beliefs $(0, \mu_1)$ *given out-of-equilibrium belief* μ_0, π' could induce a different pair of posterior beliefs for different out-of-equilibrium interim beliefs. However, note that, if deviation π' induces out-of-equilibrium belief μ_{θ_R} both types of the coalition would prefer to deviate to that experiment, as it gives a higher payoff than the payoff from any distribution with belief $\mu_s \in [\mu_1, \mu_2]$ in its support, given condition 2.14. Therefore, there always exists *some* belief such that this deviation dominates the equilibrium for both types, so the modified intuitive criterion requires that the receiver's out-of-equilibrium following that deviation be μ_0, which means that deviating to π' does indeed induce beliefs $(0, \mu_1)$.

Finally, if that deviation induces out-of-equilibrium belief μ_0, it gives a higher payoff by a concavification argument and condition 2.14.

As a result, there is always a profitable deviation from any pooling equilibrium inducing beliefs $\mu_s \in [\mu_1, \mu_2]$.

Claim 3: There is no separating equilibrium that reveals less information than the competitive equilibrium.

Proof: Let $S = \{\mu_s\}$ the support of the distribution over posterior beliefs induced by some separating equilibrium. Then this support S must include the set of beliefs $\{\mu_{s_0}(\theta_L), \mu_{s_1}(\theta_L), 1\}$, where $\mu_{s_0}(\theta_L), \mu_{s_1}(\theta_L)$ are two beliefs such that $\mu_{s_0}(\theta_L) < \mu_{\theta_L} < \mu_{s_1}(\theta_L) < 1$.

Therefore, a separating equilibrium creates a more spread-out distribution over posterior beliefs than π^c and is therefore more informative, given Lemma 8.

□

Proof of Proposition 9. The proof is constructive and provides an example of payoffs that satisfy all the conditions given in proposition 8 given the private information structure. I show that the example provided in Section 2.4.2 is a collusion-preferred pair of equilibria.

Suppose that the payoffs take the values given in (2.16) and the information structure takes the form of (2.17), then

1. Condition 2.12 is satisfied as $\frac{\mu_1}{\mu_2} = 0.43 \leq \frac{1}{2} = \frac{w_2^N}{w_3^N}$
2. Condition 2.13 is satisfied as $\frac{(1-\mu_2)(\mu_0(1-\mu_1)-\mu_{\theta_L}(\mu_0-\mu_1))}{(1-\mu_1)(\mu_0(1-\mu_2)+\mu_{\theta_L}(\mu_2-\mu_0))} = 0.832 \leq 0.833 = \frac{w_2^I}{w_1^I}$
3. Condition 2.14 is satisfied as $\frac{\mu_{\theta_L}\mu_2(1-\mu_0)}{\mu_1(\mu_{\theta_L}\mu_2(1-\mu_0)+\mu_0(1-\mu_{\theta_L})(1-\mu_2))} = 0.69 \leq 1.73 = \frac{w^c(\mu_2)}{w^c(\mu_1)}$
4. Condition 2.15 is satisfied as $\frac{\mu_1(\mu_2-\mu_0)}{\mu_0(\mu_2-\mu_1)} = 0.3 > 0.26 = \frac{w^c(\mu_2)-w_3^c}{w^c(\mu_2)-w^c(\mu_1)}$

□

Proof of Proposition 10. This proposition follows directly from Gentzkow & Kamenica (2017*b*) since the environment is Blackwell-connected here.

In particular, if neither sender has information, then the senders' interim beliefs are equal to the prior and condition 2.3 reduces to:

$$\mathbb{E}_{\mu_0}\left[V_I(\mu_1,\mu_2,\mu_0)\right] > \mathbb{E}_{\mu_0}\left[V_I(m_1,m_2,\mu_0)\right] \text{ for all } (m_1,m_2)\in[0,\mu_1)\times(\mu_2,1]$$

In addition, payoffs need to satisfy condition 2.2:

$$\mathbb{E}_{\mu_0}\left[V_N(\mu_1,\mu_2,\mu_0)\right] > \mathbb{E}_{\mu_0}\left[V_N(m_1,m_2,\mu_0)\right]$$

for all $(m_1,m_2)\in[0,\mu_1)\times(\mu_2,1]$, for collusion to be more informative than competition.

Taking a convex combination of the two inequalities with weight $\alpha\in[0,1]$ given that $(m_1,m_2)=(0,\mu_2)$ gives:

$$\mathbb{E}_{\mu_0}\left[\alpha V_I(\mu_1,\mu_2,\mu_0)+(1-\alpha)V_N(\mu_1,\mu_2,\mu_0)\right] > \mathbb{E}_{\mu_0}\left[\alpha V_I(0,\mu_2,\mu_0)+(1-\alpha)V_N(0,\mu_2,\mu_0)\right]$$
$$\Leftrightarrow \mathbb{E}_{\mu_0}\left[V_m(\mu_1,\mu_2,\mu_0)\right] > \mathbb{E}_{\mu_0}\left[V_m(0,\mu_2,\mu_0)\right]$$

Which contradicts condition 2.5 (with $\mu_\theta=\mu_0$). Therefore, conditions 2.2, 2.3, and 2.5 can never be satisfied at the same time when $\mu_{\theta_L}=\mu_{\theta_R}=\mu_0$.

□

Proof of Proposition 11. To prove that the least informative collusive equilibrium is not strictly more informative than the least-collusive competitive equilibrium when senders have full information, I show that the only (and therefore least informative) competitive equilibrium is fully revealing.

First note that if any of the senders play a separating strategy, then the receiver fully learns the state in equilibrium regardless of the experiment chosen, so we can restrict attention to pooling equilibria.

Suppose by contradiction that the two types of the two senders pool on an experiment that induces posterior beliefs such that at least one of these beliefs m satisfies $0<m<1$.

If $\mu_1<m<1$, type θ_L of Sender I can deviate to a fully revealing experiment. This increases Sender I's expected utility to w_1^I with probability 1, which is greater than the utility from any other experiment that induces some belief m satisfying $\mu_1<m<1$.

In particular, since type θ_R of Sender I would never deviate to such a experiment, the out-of-equilibrium put only weight on θ_L, so we do not need to consider how the receiver would update a degenerate belief that contradicts the fully revealing experiment.

Similarly, if $0 < m \leq \mu_1$, type θ_R of Sender N would prefer to deviate to a fully revealing experiment since this gives expected utility w_3^N with probability 1, which is greater than the utility from any other experiment that induces some belief m satisfying $\mu_1 < m < 1$.

Hence the unique competitive equilibrium must be fully revealing, and the least informative collusive equilibrium cannot possibly be strictly more informative.

$\square$

Proof of Proposition 12. If the conditions in proposition 8 are satisfied, then the least informative equilibrium in collusion is more informative than the least-information equilibrium in competition.

Therefore, if the receiver's welfare is increasing in the amount of information available, then the receiver is better-off under collusion than competition, when focusing on the least informative equilibrium. $\square$

Proof of Proposition 13. Since the uninformed sender can unilaterally deviate to an experiment inducing τ^m in competition without changing the interim beliefs of the receiver, then distribution τ^m must make the uninformed sender worse-off. Otherwise, he would deviate to it in competition, which contradicts that τ^c is an equilibrium in competition. Therefore, the uninformed sender must be (ex-ante) worse-off in collusion.

The equilibrium distribution in collusion induces posterior beliefs $(0, \mu_2)$. Therefore, type θ_R of the informed sender expects to get w_2^I with probability 1 in collusion, instead of some mixture of w_1^I and w_1^N in competition. Therefore, the payoff in collusion is higher than in competition as $w_1^I > w_1^N$. Therefore, type θ_R of the informed sender must be (ex-ante) worse-off in collusion.

Finally, suppose by contradiction that type θ_L of Sender I is worse-off under competition than under collusion. Then this type could deviate to the more informative distribution τ^m in competition. However, condition 2.13 in 8 implies that this deviation is not profitable for any out-of -equilibrium belief of the receiver. Therefore, type θ_L of Sender I is also worse-off in collusion than in competition.

Since both types of Sender I are worse-off, then Sender I must be worse-off ex-ante too.

$\square$

C Proofs of results in Chapter 3

Proof of proposition 14. For the first and second part of the proposition ($a_2 = 0$ in any equilibrium, and the voter believes $a_2 = 0$ with probability 1): for any type of the politician, and for any realisation of the random shock to utility, the politician's second-period payoff is strictly decreasing in effort. Since the politician cannot credibly commit to a second-period action in the first period, the unique optimal strategy is to exert zero effort. Common knowledge of rationality implies that the only possible belief for the voter is that the politician will exert zero effort with probability 1.

For the third part, given the optimal choice of effort of the politician in the second period, the voter's choice of $r(u_1)$ at the end of the first period maximises

$$r(u_1)\mathbb{E}[u_2(0,\theta_I)|u_1] + (1 - r(u_1))\mathbb{E}[u_2(0,\theta_C)]$$

The voter will therefore re-elect the incumbent if and only if (assuming re-election when indifferent):

$$\mathbb{E}[u_2(0,\theta_I)|u_1] \geq \mathbb{E}[u_2(0,\theta_C)]$$

Given $u_2 = f(a_2,\theta) + \varepsilon_2$, $\theta_L, \theta_H > 0$, and $\frac{\partial f(a_t,\theta)}{\partial\theta} \geq 0$, this is equivalent to

$$\mathbb{E}[\theta_I|u_1] \geq \mathbb{E}[\theta_C]$$

Finally, given the prior belief over the type of a politician in the pool of challenger, $\mathbb{P}(\theta = \theta_H) = p$, the voter re-elects the incumbent if and only if

$$\mathbb{P}(\theta_I = \theta_H|u_1) \geq p \tag{27}$$

The voter updates her beliefs using Bayes' rule, given her beliefs over the choice of effort of the politician in the first period. Let $\tilde{a}_1$ denote the (degenerate) belief of the voter over the politician's choice of effort. The voter's naivete only affects her beliefs about the noise ε and about the politician's own beliefs about that noise. Therefore, the voter correctly believes that the politician has got no private information on her type and the left-hand side of equation (27) is then equal to

$$\mathbb{P}(\theta_I = \theta_H|u_1,\tilde{a}_1) = \frac{\mathbb{P}(u_1|\theta_I = \theta_H,\tilde{a}_1)\mathbb{P}(\theta_I = \theta_H)}{\mathbb{P}(u_1|\theta_I = \theta_H,\tilde{a}_1)\mathbb{P}(\theta_I = \theta_H) + \mathbb{P}(u_1|\theta_I = \theta_L,\tilde{a}_1)\mathbb{P}(\theta_I = \theta_L)}$$

Which given the density of the voter's prior over ε_t, g_V, is:

$$\mathbb{P}(\theta_I = \theta_H | u_1 \tilde{a}_1) = \frac{p g_V(u_1 - f(\tilde{a}_1, \theta_H))}{p g_V(u_1 - f(\tilde{a}_1, \theta_H)) + (1-p) g_V(u_1 - f(\tilde{a}_1, \theta_L))}$$

Therefore, the condition on u_1 for the incumbent to be re-elected is

$$\frac{p g_V(u_1 - f(\tilde{a}_1, \theta_H)}{p g_V(u_1 - f(\tilde{a}_1, \theta_H)) + (1-p) g_V(u_1 - f(\tilde{a}_1, \theta_L))} \geq p$$

Or equivalently,

$$\frac{g_V(u_1 - f(\tilde{a}_1, \theta_H))}{g_V(u_1 - f(\tilde{a}_1, \theta_L))} \geq 1 \tag{28}$$

Since the voter's prior distribution G_V over the random shock ε satisfies the monotone likelihood ratio property (MLRP) (by the normality assumption), the left-hand side of inequality (28) is increasing in u_1. This implies that the voter re-elects the politician if and only if her first period payoffs are higher than a threshold. Let $\bar{u}$ denote this threshold.

In particular, this threshold should solve $\frac{g_V(\bar{u} - f(\tilde{a}_1, \theta_H))}{g_V(\bar{u} - f(\tilde{a}_1, \theta_L))} = 1$, or equivalently

$$g_V(\bar{u} - f(\tilde{a}_1, \theta_H)) = g_V(\bar{u} - f(\tilde{a}_1, \theta_L))$$

This concludes the proof of the third statement and of the Proposition.

□

Proof of Proposition 15. First consider the **politician's problem** given some arbitrary threshold $\bar{u}$. The effort choice is determined by the first-order condition 3.6. I express this condition in terms of the model parameters below, let a_1^* the optimal effort given an arbitrary threshold $\bar{u}$:

$$\begin{aligned} c'(a_1^*) = p\delta g\Big(\bar{u} - f(a_1^*, \theta_H)\Big) \frac{\partial f(a_1, \theta_H)}{\partial a_1}\Big|_{a_1 = a_1^*} \\ + (1-p)\delta g\Big(\bar{u} - f(a_1^*, \theta_L)\Big) \frac{\partial f(a_1, \theta_L)}{\partial a_1}\Big|_{a_1 = a_1^*} \end{aligned} \tag{29}$$

The first-order condition identifies the optimal effort if the second-order condition is satisfied. Expanding condition 3.7 gives:

$$\begin{aligned} c''(a_1^*) \geq\ & p\delta\Bigg[g\left(\bar{u}-f(a_1^*,\theta_H)\right)\frac{\partial^2 f(a_1,\theta_H)}{\partial a_1^2}\Big|_{a_1=a_1^*} \\ &\qquad - g'\left(\bar{u}-f(a_1^*,\theta_H)\right)\left(\frac{\partial f(a_1,\theta_H)}{\partial a_1}\Big|_{a_1=a_1^*}\right)^2\Bigg] \\ &+ (1-p)\delta\Bigg[g\left(\bar{u}-f(a_1^*,\theta_L)\right)\frac{\partial^2 f(a_1,\theta_L)}{\partial a_1^2}\Big|_{a_1=a_1^*} \\ &\qquad - g'\left(\bar{u}-f(a_1^*,\theta_L)\right)\left(\frac{\partial f(a_1,\theta_L)}{\partial a_1}\Big|_{a_1=a_1^*}\right)^2\Bigg] \end{aligned} \tag{30}$$

Voter's choice of threshold: given that the voter has correct beliefs over the politician's beliefs about the distribution of ε_1, and given that the politician's choice of effort is based on her own belief about that distribution and her belief about the voter's equilibrium threshold, the voter uses the solution to the first-order condition 29 to determine the threshold. This is because the voter's belief about the variance σ_V^2 does not affect her choice of threshold. Therefore, the threshold is obtained by substituting a_1^* into equation (3.9):

$$\bar{u} = \frac{f(a_1^*,\theta_H)+f(a_1^*,\theta_L)}{2}$$

Substituting this voter's re-election threshold and using assumption 4, we find that the equilibrium effort a_P^S must satisfy:

$$\begin{aligned} c'(a_S^P) = \frac{\delta}{\sigma}\phi\left(\frac{f(a_S^P,\theta_H)-f(a_S^P,\theta_L)}{2\sigma}\right) & \\ \times\left[p\frac{\partial f(a_1,\theta_H)}{\partial a_1}\Big|_{a_1=a_S^P} + (1-p)\frac{\partial f(a_1,\theta_L)}{\partial a_1}\Big|_{a_1=a_S^P}\right] & \end{aligned} \tag{31}$$

And that it is a maximum if:

$$\begin{aligned} c''(a_S^P) \geq\ & \frac{\delta}{\sigma}\Bigg[\frac{1}{2\sigma}\phi'\left(\frac{f(a_S^P,\theta_H)-f(a_S^P,\theta_L)}{2\sigma}\right)\left(\frac{\partial f(a_S^P,\theta_H)}{\partial a_1}-\frac{\partial f(a_S^P,\theta_L)}{\partial a_1}\right) \\ &\quad \times\left(p\frac{\partial f(a_S^P,\theta_H)}{\partial a_1}+(1-p)\frac{\partial f(a_S^P,\theta_L)}{\partial a_1}\right) \\ &\quad + \phi\left(\frac{f(a_S^P,\theta_H)-f(a_S^P,\theta_L)}{2\sigma}\right)\left[p\frac{\partial^2 f(a_S^P,\theta_H)}{\partial a_1^2}+(1-p)\frac{\partial^2 f(a_S^P,\theta_L)}{\partial a_1^2}\right]\Bigg] \end{aligned}$$

This proves the first and second statement of the proposition.

To prove existence, note that the left-hand side of equation (31) is strictly increasing in a and is equal to 0 at $a = 0$ and goes to infinity as a tends to infinity, given the properties of c in assumption 5. The right-hand side of (31), denoted $B(a)$ satisfies

$$\lim_{a\to 0} B(a) > 0$$

Since $\lim_{a\to 0} \frac{\delta}{\sigma}\phi\left(\frac{f(a,\theta_H)-f(a,\theta_L)}{2\sigma}\right)$ is strictly positive given $\left|\lim_{a\to 0} f(a,\theta)\right| < +\infty$, and $\lim_{a\to 0}\left[p\frac{\partial f(a,\theta_H)}{\partial a} + (1-p)\frac{\partial f(a,\theta_L)}{\partial a}\right] > 0$ since $\lim_{a\to 0}\frac{\partial f(a,\theta)}{\partial a} > 0$, given assumption 3.

In addition, $B(a)$ satisfies,

$$\lim_{a\to +\infty} B(a) < +\infty$$

Since $\lim_{a\to +\infty} \frac{\delta}{\sigma}\phi\left(\frac{f(a,\theta_H)-f(a,\theta_L)}{2\sigma}\right)$ is finite given $\sigma > 0$, and since

$$\lim_{a\to +\infty}\left[p\frac{\partial f(a,\theta_H)}{\partial a} + (1-p)\frac{\partial f(a,\theta_L)}{\partial a}\right] < +\infty$$

given $\lim_{a\to 0}\frac{\partial f(a,\theta)}{\partial a} < +\infty$ as stated in assumption 3.

Therefore, we have that

$$c'(0) = 0 < \lim_{a\to 0} B(a) \quad \text{and} \quad \lim_{a\to +\infty} B(a) < \lim_{a\to +\infty} c'(a) = +\infty$$

So by the intermediate value theorem, there exists $a^* \in (0, +\infty)$ such that the left-hand side of equation (31) equals the right-hand side if $a = a^*$. The strategy profile $(a, \bar{u}) = \left(a^*, \frac{f(a^*,\theta_H)+f(a^*,\theta_L)}{2}\right)$ is then an equilibrium.

For uniqueness, a sufficient condition is that the function

$$T(a) := \delta\left[1 - p\Phi\left(\frac{f(a,\theta_L) - f(a,\theta_H)}{2\sigma}\right) - (1-p)\Phi\left(f(a,\theta_H) - (f(a,\theta_L))\right)\right] - c(a)$$

is strictly concave. In that case, $T'(a) < 0$ for any a, so given the existence result above, $T(a) = 0$ for a unique value of a, and $\bar{u}$ is then also uniquely determined. Note that the concavity of $T(a)$ also implies that the second-order condition is satisfied. □

Proof of proposition 16. The proof of the first three statements in the proposition is given in the text. For the last statement, note that since the voter believes the incumbent chooses effort a^V, and the incumbent is aware of the voter's bias and naivety, then the incumbent also believes that the voter believes she chooses effort a^V. Similarly, since the incumbent believes the voter will re-elect if and only if her payoff is above the threshold $\bar{u}^V$

, then the voter believes that the incumbent believes the voter uses this threshold. Therefore, both player's beliefs over the other player's beliefs over the equilibrium strategies are correct, and as a result any higher-order beliefs must be correct too.

□

Proof of proposition 17. The existence and uniqueness of a RCE requires that both the incumbent's action a^I and the voter's beliefs over the incumbent's action a^V exist and are unique.

The voter's belief about the incumbent's action a^V solves:

$$c'(a^V) = \delta \frac{1}{\sigma_V} \phi \left(\frac{f(a^V, \theta_H) - f(a^V, \theta_L)}{2\sigma_V} \right) \left[p \frac{\partial f(a_1, \theta_H)}{\partial a_1}\Big|_{a_1 = a^V} + (1-p) \frac{\partial f(a_1, \theta_L)}{\partial a_1}\Big|_{a_1 = a^V} \right] \tag{32}$$

This is the same equation as equation (31) and therefore the same arguments as in the proof of Proposition 15 ensure the existence and the uniqueness of a^V.

Given a unique a^V, the incumbent's equilibrium action is defined by equation (3.14) reproduced below:

$$\begin{aligned} c'(a^I) = p\delta \frac{1}{\sigma} \phi \left(\frac{\frac{f(a^V, \theta_H) + f(a^V, \theta_L)}{2} - f(a^I, \theta_H)}{\sigma} \right) \frac{\partial f(a_1, \theta_H)}{\partial a_1}\Big|_{a_1 = a^I} \\ + (1-p)\delta \frac{1}{\sigma} \phi \left(\frac{\frac{f(a^V, \theta_H) + f(a^V, \theta_L)}{2} - f(a^I, \theta_L)}{\sigma} \right) \frac{\partial f(a_1, \theta_L)}{\partial a_1}\Big|_{a_1 = a^I} \end{aligned}$$

The left-hand side is 0 at $a^I = 0$ and tends to infinity as $a^I \to +\infty$ given assumption 5. The right-hand side is strictly positive as $a^I \to 0$ since

1. $\phi \left(\frac{\frac{f(a^V, \theta_H) + f(a^V, \theta_L)}{2} - f(a^I, \theta)}{\sigma} \right) > 0$ for any a^V, any $\theta \in \{\theta_L, \theta_H\}$, and any a^I as $\left| \lim_{a \to 0} f(a, \theta) \right| < +\infty$,

2. and $\lim_{a \to 0} \frac{\partial f(a, \theta)}{\partial a} > 0$ given assumption 3.

Finally, the limit of the right-hand side is less than infinity as $a^I \to +\infty$ since

$$\lim_{a^I \to +\infty} \phi \left(\frac{\frac{f(a^V, \theta_H) + f(a^V, \theta_L)}{2} - f(a^I, \theta_H)}{\sigma} \right) < +\infty$$

as $\sigma > 0$ and since $\lim_{a \to +\infty} \frac{\partial f(a, \theta)}{\partial a} < +\infty$ given assumption 3. Therefore, by the intermediate value theorem there exists a^I such that equation (3.14) is satisfied.

A sufficient condition for uniqueness is that the derivative of

$$V^{RCE} = \frac{\delta}{\sigma}\left[p\Phi\left(\frac{u - f(a,\theta_H)}{\sigma}\right) + (1-p)\Phi\left(\frac{u - f(a,\theta_L)}{\sigma}\right)\right]$$

is strictly negative for $u = \frac{f(a^V,\theta_H)+f(a^V,\theta_L)}{2}$. Indeed,

$$\begin{aligned}c'(a^I) = p\delta\frac{1}{\sigma}\phi\left(\frac{\frac{f(a^V,\theta_H)+f(a^V,\theta_L)}{2} - f(a^I,\theta_H)}{\sigma}\right)\frac{\partial f(a,\theta_H)}{\partial a}\Big|_{a=a^I} \\ + (1-p)\delta\frac{1}{\sigma}\phi\left(\frac{\frac{f(a^V,\theta_H)+f(a^V,\theta_L)}{2} - f(a^I,\theta_L)}{\sigma}\right)\frac{\partial f(a^I,\theta_L)}{\partial a}\Big|_{a=a^I}\end{aligned}$$

Then for all $a < a^I$,

$$\begin{aligned}c'(a) < p\delta\frac{1}{\sigma}\phi\left(\frac{\frac{f(a^V,\theta_H)+f(a^V,\theta_L)}{2} - f(a,\theta_H)}{\sigma}\right)\frac{\partial f(a,\theta_H)}{\partial a} \\ + (1-p)\delta\frac{1}{\sigma}\phi\left(\frac{\frac{f(a^V,\theta_H)+f(a^V,\theta_L)}{2} - f(a,\theta_L)}{\sigma}\right)\frac{\partial f(a,\theta_L)}{\partial a}\end{aligned}$$

And for all $a > a^I$,

$$\begin{aligned}c'(a) > p\delta\frac{1}{\sigma}\phi\left(\frac{\frac{f(a^V,\theta_H)+f(a^V,\theta_L)}{2} - f(a,\theta_H)}{\sigma}\right)\frac{\partial f(a,\theta_H)}{\partial a} \\ + (1-p)\delta\frac{1}{\sigma}\phi\left(\frac{\frac{f(a^V,\theta_H)+f(a^V,\theta_L)}{2} - f(a,\theta_L)}{\sigma}\right)\frac{\partial f(a,\theta_L)}{\partial a}\end{aligned}$$

So that the two sides only intersect once. □

Proof of lemma 4. Recall that the different optimal levels of efforts solve the following equations:

The voter's belief about equilibrium effort solves

$$c'(a^V) = \frac{\delta}{\sigma^V}\phi\left(\frac{\theta_H}{2\sigma_V}\right)$$

The politician's choice of effort in the absence of overconfident voter solves

$$c'(a^R) = \delta\phi\left(\frac{\theta_H}{2}\right)$$

The politician's choice of effort in the presence of overconfident voters solves

$$c'(a^I) = \delta\left[p\phi\left(\frac{2(a^V - a^I) - \theta_H}{2}\right) + (1-p)\phi\left(\frac{2(a^V - a^I) + \theta_H}{2}\right)\right]$$

Since assumption 6 implies that assumption 3 is satisfied, then given assumption 5 and Propositions 15 and 17 we can conclude that a^V and a^R exist and are unique, and that a^I exists.

For the uniqueness of a^I, I show that $\delta\left[p\phi\left(\frac{2(a^V-a)-\theta_H}{2}\right) + (1-p)\phi\left(\frac{2(a^V-a)+\theta_H}{2}\right)\right]$ is decreasing if and only if $a > \bar{a}$ where $\bar{a}$ solves

$$(1-p)\left(\frac{2a^V + \theta_H}{2} - \bar{a}\right)\exp\left(2\theta_H(a^V - \bar{a})\right) + p\left(\frac{2a^V - \theta_H}{2} - \bar{a}\right) = 0$$

Define the right-hand side of equation (3.14) under assumption 6 as

$$B(a) \equiv \delta\left[p\phi\left(\frac{2(a^V - a) - \theta_H}{2}\right) + (1-p)\phi\left(\frac{2(a^V - a) + \theta_H}{2}\right)\right]$$

Then,

$$\begin{aligned}
& B'(a) \leq 0 \\
\Leftrightarrow\quad & \frac{\partial}{\partial a}\left[\delta\left[p\phi\left(\frac{2(a^V - a) - \theta_H}{2}\right) + (1-p)\phi\left(\frac{2(a^V - a) + \theta_H}{2}\right)\right]\right] \leq 0 \\
\Leftrightarrow\quad & p\left(\frac{2(a^V - a) - \theta_H}{2}\right)\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{(2(a^V - a) - \theta_H)^2}{8}\right) \\
& \quad + (1-p)\left(\frac{2(a^V - a) + \theta_H}{2}\right)\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{(2(a^V - a) + \theta_H)^2}{8}\right) \leq 0 \\
\Leftrightarrow\quad & p\left(\frac{2a^V - \theta_H}{2} - a\right) + (1-p)\frac{\exp\left(-\frac{(2(a^V-a)+\theta_H)^2}{8}\right)}{\exp\left(\frac{(2(a^V-a)+\theta_H)^2}{2}\right)} \leq 0 \\
\Leftrightarrow\quad & (1-p)\left(\frac{2a^V + \theta_H}{2} - a\right)\exp\left(2\theta_H(a^V - a)\right) + p\left(\frac{2a^V - \theta_H}{2} - a\right) \leq 0
\end{aligned}$$

In addition, $B(a)$ only equals 0 once. Suppose that for some a, $B(a) \leq 0$.

1. Then at least one of $\frac{2a^V+\theta_H}{2} - a \leq 0$ or $\frac{2a^V-\theta_H}{2} - a$ holds.

2. Since $\frac{2a^V-\theta_H}{2} < \frac{2a^V+\theta_H}{2}$, then at least $\frac{2a^V-\theta_H}{2} - a$ must hold.

3. If $\frac{2a^V+\theta_H}{2} - a \leq 0$, then for any $a' > a$, $\frac{2a^V-\theta_H}{2} - a' \leq 0$, $\frac{2a^V+\theta_H}{2} - a' \leq 0$ and therefore $(1-p)\left(\frac{2a^V+\theta_H}{2} - a'\right)\exp\left(2\theta_H(a^V - a')\right) + p\left(\frac{2a^V-\theta_H}{2} - a\right) \leq 0$.

4. If $\frac{2a^V+\theta_H}{2} - a \geq 0$, then

$$\begin{aligned} &\frac{\partial}{\partial a}(1-p)\left(\frac{2a^V+\theta_H}{2} - a\right)\exp\left(2\theta_H(a^V-a)\right) \\ &\qquad = -(1-p)\exp\left(2\theta_H(a^V-a)\right)\left(1+2\theta_H\left(\frac{2a^V+\theta_H}{2}-a\right)\right) \\ &\qquad < 0 \end{aligned}$$

So for any $a' > a$ such that $a' \leq \frac{2a^V+\theta_H}{2} - a$, we also have $B(a') \leq 0$.

Finally, note that $a = 0$,

$$\begin{aligned} B(0) &= \delta\left[(1-p)\left(\frac{2a^V+\theta_H}{2}-a\right)\exp\left(2\theta_H(a^V-a)\right) + p\left(\frac{2a^V-\theta_H}{2}-a\right)\right] \\ &\geq 0 \end{aligned}$$

And as $a \to \infty$,

$$\begin{aligned} \lim_{a\to\infty} B(a) &= \lim_{a\to\infty}\delta\left[(1-p)\left(\frac{2a^V+\theta_H}{2}-a\right)\exp\left(2\theta_H(a^V-a)\right) + p\left(\frac{2a^V-\theta_H}{2}-a\right)\right] \\ &\leq 0 \end{aligned}$$

Therefore, there exists $\bar{a}$ such that

$$B(a) \leq 0 \quad \Leftrightarrow \quad a \geq \bar{a}$$

□

Proof of Proposition 18. Given lemma 4, the proof of proposition 18 proceeds as follows. First, I show that a^I is higher than a^R if and only if $a^V > a^R$. Then I show that $a^V > a^R$ if and only if $\sigma_V > s^*(\theta)$. Finally, I show that $s^*(\theta) \in [0,1]$ and that $s^*(\theta)$ is decreasing in θ_H.

Lemma 10. *Suppose that assumptions 6 is satisfied, and $\bar{a} \leq 0$, then $a^I > a^R$ if and only if $a^V > a^R$.*

Proof of lemma 10. First note the following.

At $a = a^V$,

$$\begin{aligned} B(a^V) &= \delta \left[p\phi \left(\frac{2(a^V - a^V) - \theta_H}{2} \right) + (1-p)\phi \left(\frac{2(a^V - a^V) + \theta_H}{2} \right) \right] \\ &= \delta\phi \left(\frac{\theta_H}{2} \right) \\ &= c'(a^R) \end{aligned}$$

Given that $c'(a)$ is increasing in a and $B(a)$ is decreasing in a (as $\bar{a} \leq 0$), we have:

1. $a^V > a^R \Rightarrow c'(a^V) > c'(a^R)$
2. $a^V > a^R \Rightarrow B(a^V) < B(a^R)$

Thus,

$$a^V > a^R \quad \Rightarrow \quad c'(a^V) > c'(a^R) = B(a^V) \text{ and } c'(a^R) = B(a^V) < B(a^R)$$

Therefore, since $B(a)$ and $c'(a)$ are continuous, there exists a such that $a^R < a < a^V$ and $c'(a) = B(a)$. Since a^I solves $c'(a^I) = B(a^I)$ and is unique, then

$$a^V > a^R \quad \Rightarrow \quad a^R < a^I < a^V$$

Similarly, $a^V < a^R \Rightarrow a^V < a^I < a^R$, which concludes the proof.

□

I then prove the following lemma:

Lemma 11. *There exists a threshold $s^*(\theta_H)$ such that $a^V < a^R$ if and only $\sigma_V^2 < s^*(\theta_H)$.*

Proof of lemma 11. Given that $c'(.)$ is increasing, $a^V < a^R$ if and only if $c'(a^V) < c'(a^R)$, and given equations (3.15) and (3.12) this holds if and only if $\frac{\delta}{\sigma^V}\phi\left(\frac{\theta_H}{2\sigma_V}\right) < \delta\phi\left(\frac{\theta_H}{2}\right)$.

Or equivalently, if and only if

$$\begin{aligned} \frac{1}{\sigma_V} &< \frac{\exp\left(-\frac{1}{2}\left(\frac{\theta_H}{2}\right)^2\right)}{\exp\left(-\frac{1}{2}\left(\frac{\theta_H}{2\sigma_V}\right)^2\right)} \\ \Leftrightarrow \quad \frac{1}{\sigma_V} &< \exp\left(\frac{1}{2}\left[\left(\frac{\theta_H}{2\sigma_V}\right)^2 - \left(\frac{\theta_H}{2}\right)^2\right]\right) \\ \Leftrightarrow \quad -\frac{\theta_H^2}{8} &< \ln(\sigma_V)\left(\frac{\sigma_V^2}{1-\sigma_V^2}\right) \end{aligned}$$

The right-hand side of this inequality is decreasing in σ_V, tends to $-\frac{1}{2}$ as $\sigma_V \to 1$ and to 0 as $\sigma_V \to 0$, as shown below.

Claim 1: $\ln(\sigma_V)\left(\frac{\sigma_V^2}{1-\sigma_V^2}\right)$ is decreasing in σ_V.

Proof: First note that

$$\frac{\partial\left[\ln(x)\frac{x^2}{1-x^2}\right]}{\partial x} = \frac{x}{1-x^2} + \ln(x)\frac{2x}{(1-x^2)^2}$$

And secondly, note that

$$\frac{x}{1-x^2} + \ln(x)\frac{2x}{(1-x^2)^2} < 0 \Leftrightarrow 1 < -2\ln(x) + x^2$$

In addition, $-2\ln(x) + x^2$ is decreasing in x if $x \leq 1$ since $\frac{\partial -2\ln(x)+x^2}{\partial x} = -\frac{2}{x} + 2x$ and at $x = 1$, $-2\ln(x) + x^2 = 1$, so that $-2\ln(x) + x^2 \geq 1$ for any $x \in [0,1]$.

Therefore, $\ln(\sigma_V)\left(\frac{\sigma_V^2}{1-\sigma_V^2}\right)$ is decreasing in σ_V.

Claim 2: $\lim_{\sigma_V \to 0} \ln(\sigma_V)\left(\frac{\sigma_V^2}{1-\sigma_V^2}\right) = 0$.

Proof: Using L'Hôpital's rule

$$\lim_{\sigma_V \to 0} \ln(\sigma_V)\left(\frac{\sigma_V^2}{1-\sigma_V^2}\right) = \lim_{\sigma_V \to 0} \frac{\ln(\sigma_V)}{\left(\frac{1-\sigma_V^2}{\sigma_V^2}\right)} = \lim_{\sigma_V \to 0} \frac{\left(\frac{1}{\sigma_V}\right)}{\left(-\frac{2}{\sigma_V^3}\right)} = \lim_{\sigma_V \to 0} -\frac{\sigma_V^2}{2} = 0$$

Claim 3: $\lim_{\sigma_V \to 1} \ln(\sigma_V)\left(\frac{\sigma_V^2}{1-\sigma_V^2}\right) = -\frac{1}{2}$.

Proof: Using L'Hôpital's rule

$$\lim_{\sigma_V \to 1} \ln(\sigma_V)\left(\frac{\sigma_V^2}{1-\sigma_V^2}\right) = \lim_{\sigma_V \to 1} \frac{\ln(\sigma_V)}{\left(\frac{1-\sigma_V^2}{\sigma_V^2}\right)} = \lim_{\sigma_V \to 1} \frac{\left(\frac{1}{\sigma_V}\right)}{\left(-\frac{2}{\sigma_V^3}\right)} = \lim_{\sigma_V \to 1} -\frac{\sigma_V^2}{2} = -\frac{1}{2}$$

As a result, combining claims 1, 2 and 3, we get that there exists a threshold $s^*(\theta_H)$ such that $-\frac{\theta_H^2}{8} < \ln(\sigma_V)\left(\frac{\sigma_V^2}{1-\sigma_V^2}\right)$ if and only if $\sigma_V < s^*(\theta_H)$.

□

The first part of Proposition 18 then follows directly from lemmas 10 and 11.

I now show that $s^*(\theta_H) \in [0,1]$ provided that $\theta_H < 2$. By claims 2 and 3 and the intermediate value theorem, the equation $-\frac{\theta_H^2}{8} = \ln(\sigma_V)\left(\frac{\sigma_V^2}{1-\sigma_V^2}\right)$ has a unique solution in $[0,1]$ if $\frac{1}{2} < -\frac{\theta_H^2}{8} < 0$.

As $\theta_H > 0$, then $-\frac{\theta_H^2}{8} < 0$ is always satisfied. In addition, $\frac{1}{2} < -\frac{\theta_H^2}{8}$ if and only if $\theta_H < 2$.

Finally, as θ_H increases, $-\frac{\theta_H^2}{8}$ decreases, so since $\ln(\sigma_V)\left(\frac{\sigma_V^2}{1-\sigma_V^2}\right)$ is decreasing, then $s^*(\theta_H)$ is increasing in θ_H.

$\square$

Proof of proposition 19. First note that the condition for selection under overconfidence to be higher than for rational voters is derived as follows

$$\begin{aligned}
S(\bar{u}) &= p\left[\mathbb{P}(u > \bar{u} \mid \theta_H)\theta_H + \mathbb{P}(u < \bar{u} \mid \theta_H)(p\theta_H + (1-p)\theta_L)\right] \\
&\quad + (1-p)\left[\mathbb{P}(u > \bar{u} \mid \theta_L)\theta_L + \mathbb{P}(u < \bar{u} \mid \theta_L)(p\theta_H + (1-p)\theta_L)\right] \\
&= p(1-p)(\theta_H - \theta_L)\left[\mathbb{P}(u > \bar{u} \mid \theta_H) - \mathbb{P}(u > \bar{u} \mid \theta_L)\right] + p\theta_H + (1-p)\theta_L
\end{aligned}$$

And in particular,

$$\mathbb{P}(u > \bar{u}^V \mid \theta_H) - \mathbb{P}(u > \bar{u}^V \mid \theta_L) > \mathbb{P}(u > \bar{u}^R \mid \theta_H) - \mathbb{P}(u > \bar{u}^R \mid \theta_L)$$

$$\Leftrightarrow$$

$$\begin{aligned}
\Phi\left(\frac{f(a^V,\theta_H) + f(a^V,\theta_L) - 2f(a^I,\theta_L)}{2\sigma}\right) - \Phi\left(\frac{f(a^V,\theta_H) + f(a^V,\theta_L) - 2f(a^I,\theta_H)}{2\sigma}\right) \\
> \Phi\left(\frac{f(a^R,\theta_H) - f(a^R,\theta_L)}{2\sigma}\right) - \Phi\left(\frac{f(a^R,\theta_L) - f(a^R,\theta_H)}{2\sigma}\right)
\end{aligned}$$

Which under assumptions 5 and 6, and given $\theta_L = 0$ and $\sigma = 1$ reduces to

$$\Phi\left(\frac{2(a^V - a^I) + \theta_H}{2}\right) - \Phi\left(\frac{2(a^V - a^I) - \theta_H}{2}\right) > \Phi\left(\frac{\theta_H}{2}\right) - \Phi\left(\frac{-\theta_H}{2}\right)$$

Let $A = 2(a^V - a^I)$. Then,

$$\Phi\left(\frac{\theta_H + 2(a^V - a^I)}{2}\right) - \Phi\left(\frac{-\theta_H + 2(a^V - a^I)}{2}\right) = \int_{-\theta_H + A}^{\theta_H + A} \phi\left(\frac{x}{2}\right) \mathrm{d}x$$

And

$$\int_{-\theta_H + A}^{\theta_H + A} \phi\left(\frac{x}{2}\right) \mathrm{d}x = \int_{-\theta_H}^{\theta_H} \phi\left(\frac{x}{2}\right) \mathrm{d}x - \int_{-\theta_H}^{-\theta_H + A} \phi\left(\frac{x}{2}\right) \mathrm{d}x + \int_{\theta_H}^{\theta_H + A} \phi\left(\frac{x}{2}\right) \mathrm{d}x$$

If $a^V \geq a^I$, then $A \geq 0$, so

$$\begin{aligned}
\int_{-\theta_H}^{\theta_H} \phi\left(\frac{x}{2}\right) \mathrm{d}x &= \int_{-\theta_H + A}^{\theta_H + A} \phi\left(\frac{x}{2}\right) \mathrm{d}x + \left[\int_{-\theta_H}^{-\theta_H + A} \phi\left(\frac{x}{2}\right) \mathrm{d}x - \int_{\theta_H}^{\theta_H + A} \phi\left(\frac{x}{2}\right) \mathrm{d}x\right] \\
&= \int_{-\theta_H + A}^{\theta_H + A} \phi\left(\frac{x}{2}\right) \mathrm{d}x + \left[\int_{\theta_H - A}^{\theta_H} \phi\left(\frac{x}{2}\right) \mathrm{d}x - \int_{\theta_H}^{\theta_H + A} \phi\left(\frac{x}{2}\right) \mathrm{d}x\right]
\end{aligned}$$

Where the second line follows from the symmetry of the standard normal distribution around 0. In addition, since $\theta_H > 0$, we have

$$\int_{\theta_H - A}^{\theta_H} \phi\left(\frac{x}{2}\right) \mathrm{d}x - \int_{\theta_H}^{\theta_H + A} \phi\left(\frac{x}{2}\right) \mathrm{d}x \geq 0$$

So that,

$$\int_{-\theta_H}^{\theta_H} \phi\left(\frac{x}{2}\right) \mathrm{d}x \geq \int_{-\theta_H + A}^{\theta_H + A} \phi\left(\frac{x}{2}\right) \mathrm{d}x$$

That is,

$$\Phi\left(\frac{\theta_H}{2}\right) - \Phi\left(\frac{-\theta_H}{2}\right) \geq \Phi\left(\frac{2(a^V - a^I) + \theta_H}{2}\right) - \Phi\left(\frac{2(a^V - a^I) - \theta_H}{2}\right)$$

So selection is better without overconfidence.

Similarly, if $A \leq 0$,

$$\int_{-\theta_H}^{\theta_H} \phi\left(\frac{x}{2}\right) \mathrm{d}x = \int_{-\theta_H + A}^{\theta_H + A} \phi\left(\frac{x}{2}\right) \mathrm{d}x + \left[\int_{\theta_H + A}^{\theta_H} \phi\left(\frac{x}{2}\right) \mathrm{d}x - \int_{\theta_H}^{\theta_H - A} \phi\left(\frac{x}{2}\right) \mathrm{d}x\right]$$

And $\int_{\theta_H + A}^{\theta_H} \phi\left(\frac{x}{2}\right) \mathrm{d}x - \int_{\theta_H}^{\theta_H - A} \phi\left(\frac{x}{2}\right) \mathrm{d}x \geq 0$, so that

$$\Phi\left(\frac{\theta_H}{2}\right) - \Phi\left(\frac{-\theta_H}{2}\right) \geq \Phi\left(\frac{2(a^V - a^I) + \theta_H}{2}\right) - \Phi\left(\frac{2(a^V - a^I) - \theta_H}{2}\right)$$

And selection is also better without overconfidence. □

Proof of proposition 20. Proposition 20 follows directly from the following facts:

1. The voter's expected utility is increasing in a_1: follows from assumption 3.

2. First period effort is higher under overconfidence when $\bar{a} \leq 0$ if $\sigma < s^*(\theta_H)$: follows from Proposition 18.

3. The voter's expected utility is increasing in the expected competence of the politician in office in the second period: follows from assumption 3.

4. Second period expected competence is always worse under overconfidence: follows from Proposition 19.

The condition below is necessary and sufficient for welfare to be higher under overconfidence:

$$a^I - a^R > p(1-p)\delta\theta_H \left[2\Phi\left(\frac{\theta_H}{2}\right) - \Phi\left(\frac{\theta_H - 2(a^V - a^I)}{2}\right) - \Phi\left(\frac{\theta_H + 2(a^V - a^I)}{2}\right)\right] \tag{33}$$

Indeed, the voter's expected utility when overconfident is:

$$\begin{aligned}
U^V &= p\left[\theta_H + a^I + \delta\left(\left(1 - \Phi\left(\frac{2(a^V - a^I) - \theta_H}{2}\right)\right)\theta_H + \Phi\left(\frac{2(a^V - a^I) - \theta_H}{2}\right)p\theta_H\right)\right] \\
&\quad + (1-p)\left[a^I + \delta\Phi\left(\frac{2(a^V - a^I) + \theta_H}{2}\right)p\theta_H\right] \\
&= p\theta_H + a^I + p\delta\theta_H\left[1 - (1-p)\left(\Phi\left(\frac{2(a^V - a^I) - \theta_H}{2}\right) - \Phi\left(\frac{2(a^V - a^I) + \theta_H}{2}\right)\right)\right]
\end{aligned}$$

The voter's expected utility when rational is:

$$\begin{aligned}
U^R &= p\left[\theta_H + a^R + \delta\left(\left(1 - \Phi\left(\frac{-\theta_H}{2}\right)\right)\theta_H + \Phi\left(\frac{-\theta_H}{2}\right)p\theta_H\right)\right] \\
&\quad + (1-p)\left[a^R + \delta\Phi\left(\frac{\theta_H}{2}\right)p\theta_H\right] \\
&= p\theta_H + a^R + p\delta\theta_H\left[p + (1-p)2\Phi\left(\frac{\theta_H}{2}\right)\right]
\end{aligned}$$

Comparing the two and re-arranging gives condition (33).

This condition is satisfied for instance with the cost function $c(a) = \frac{a^2}{2}$ and the following parameters:

Parameter	Value
δ	$\frac{1}{4}$
σ_V^2	$\frac{1}{4}$
θ_H	1
p	$\frac{3}{4}$

Since then $a^V = \frac{\left(\frac{1}{2}\right)}{\left(\frac{1}{2}\right)}\phi\left(\frac{1}{2\left(\frac{1}{2}\right)}\right) \simeq 0.12$, $a^R = \left(\frac{1}{2}\right)\phi\left(\frac{1}{2}\right) \simeq 0.088$, $a^I \simeq 0.089$, so $U^V = 1.044$ and $U^R = 0.997$.

□

D Formal definition of the rationalisable conjectural equilibrium

I follow the notation and structure used in Esponda (2013). Consider

1. A state space $\Omega = \{A, \bar{U}, \mathcal{G}, \mathcal{P}\}$, where $A = \mathbb{R}^2$ is the action space of the incumbent, $\bar{U} \subseteq \mathbb{R}$ is the action space of the voter (redefined in terms of the choice of a threshold strategy), $\mathcal{G}$ is the set of possible probability measures over $\varepsilon \in \mathbb{R}$, $\mathcal{P}$ the set of probability measures over the set $\{\theta_L, \theta_H\}$.

2. A belief space $\mathcal{B} = \langle \Omega, T, \xi, \lambda_V, \lambda_I \rangle$, where Ω is the state space (set of strategy profiles and fundamentals of the game) defined above, T is a set of epistemic states, $\xi : T \to \Omega$ is a mapping from epistemic states to states of nature, λ_V and λ_I are probability measures over T, representing the players' beliefs over epistemic states (and thus over the state space).

3. Two feedback partitions P_V for the voter and P_I for the incumbent politician.

Following Esponda (2013), a strategy profile is a rationalisable conjectural equilibrium if there exists a belief space $\mathcal{B}$ such that players maximise their expected utility, given their beliefs (rationality), and these beliefs are consistent given a partition (consistency), and such that there is common knowledge of rationality and consistency.

To find the set of such equilibria, I first define the exogenous feedback partition that each player is facing, based on the timing and information structure of the model described in section 3.2. I then find the set of strategy profiles that satisfy rationality and consistency, and finally, I refine this set of equilibria by imposing restrictions on higher-order beliefs, in particular, common knowledge of rationality and consistency.

Given that the voter only observes her utility at the end of the first period, and at the end of the second period, I restrict the support of the beliefs over the epistemic state that generates the incumbent's actions and the distribution of the utility shock ε to be consistent with the observed utility. Given that, both the incumbent's action and the random shock have support over the entire real line, this restriction is not very stringent: any observed utility is consistent with some combination of incumbent effort and some realisation of the random shock. In addition, I make the very natural restriction that the voter should know her own action (choice of re-election threshold), and I assume that the voter knows the correct distribution of the politician's type for certain (in order to focus on overconfidence over the random utility shock). Formally, $\forall\, \omega \in \Omega$, let

$$P_V(\omega) = \{\omega' \in \Omega \mid \bar{u}(\omega') = \bar{u}(\omega),\ u_1(\omega') \in \text{supp}_\omega(\mathbb{E}_\theta[f(a_1, \theta) + \varepsilon]),\ p(\omega') = p(\omega)\} \quad (34)$$

I restrict the incumbent to know with certainty the 'correct' (objective) distribution of both the random shock, and of her own type. The incumbent also observes her own actions, and observes whether or not she is re-elected. In this case too, this assumption does not place strong restrictions on the beliefs of the incumbent, since re-election could be consistent with an infinity of combinations of re-election thresholds and realisation of the random shock. Formally, $\forall\ \omega \in \Omega$, let

$$P_I(\omega) = \{\omega' \in \Omega \mid (a_1(\omega'), a_2(\omega')) = (a_1(\omega), a_2(\omega)),\ G(\omega') = G(\omega),\ p(\omega') = p(\omega)\} \quad (35)$$

Then a strategy profile $(\bar{u}_V, (a_1^I, a_2^I))$ is an RCE if there exists $\mathcal{B} = \langle \Omega, T, \xi, \lambda_V, \lambda_I \rangle$ such that

1. *Voter's rationality*

$$\bar{u}_V \in \arg\max_{\bar{u}' \in \bar{U}} \int_{\xi(T)} u_V(\bar{u}', (a_1(\omega), a_2(\omega)), G(\omega), p(\omega)) d\delta_V \quad (36)$$

   Where δ_V is the marginal probability over strategies of the incumbent and fundamentals of the game (distribution of ε and of θ) induced by the belief λ_V in $\mathcal{B}$.

2. *Incumbent's rationality*

$$(a_1^I, a_2^I) \in \arg\max_{(a_1, a_2) \in A} \int_{\xi(T)} u_I(\bar{u}(\omega), (a_1^I, a_2^I), G(\omega), p(\omega)) d\delta_I \quad (37)$$

   Where δ_I is the marginal probability over strategies of the voter and fundamentals of the game (distribution of ε and of θ) induced by the belief λ_I in $\mathcal{B}$.

3. *Consistency of voter's beliefs* The voter must put a zero-probability weight on any epistemic state that generates a state ω that does not belong to the same partition as the state generated by the epistemic state believed by the voter.

$$\lambda_V(t)(\xi^{-1}(P_V(\xi(t))) = 1 \quad (38)$$

4. *Consistency of incumbent's beliefs* Similarly for the incumbent, let

$$\lambda_I(t)(\xi^{-1}(P_I(\xi(t))) = 1 \quad (39)$$

5. *Voter's knowledge of incumbent's rationality* Let $R_I^{\mathcal{B}} \subseteq T$ the event '*the incumbent is rational*' as defined in point 2. above. Then the voter's second-order beliefs need to satisfy, for each epistemic state t on which the voter puts positive probability $\lambda_V(t)[R_I^{\mathcal{B}}] = 1$.

6. *Incumbent's knowledge of voter's rationality* Similarly for the incumbent , let $\lambda_I(t)[R_V^{\mathcal{B}}] = 1$.

7. *Voter's knowledge of incumbent's consistency* Let $C_I^{\mathcal{B}} \subseteq T$ the event '*the incumbent has consistent beliefs*' as defined in point 4. above. Then the voter's second-order beliefs need to satisfy, for each epistemic state t on which the voter puts positive probability $\lambda_V(t)[C_I^{\mathcal{B}}] = 1$.

8. *Incumbent's knowledge of voter's consistency* Similarly for the incumbent, let $\lambda_I(t)[C_V^{\mathcal{B}}] = 1$

9. In addition, the equilibrium needs to satisfy common knowledge of these last four conditions, and of all higher-order beliefs about this common knowledge.

Without any restrictions on beliefs, the set of RCEs, as just defined, would be very large. Indeed, there are infinitely many possible distributions of the random utility shock, and infinitely many possible beliefs over these distributions, each of which is associated with a different equilibrium action of the politician and equilibrium threshold of the voter.

Therefore, I focus on the case where the voter has degenerate beliefs over (the subset of epistemic states which generates) a given distribution of the random shock, denoted $G_V(\varepsilon) \in \mathcal{G}$, while the incumbent has degenerate beliefs over (the subset of epistemic states which generates) the true (objective) distribution of the random shock, denoted $G(\varepsilon) \in \mathcal{G}$. I also focus on the case where both players have degenerate beliefs over the true distribution of the politician's type $\mathbb{P}(\theta = \theta_H) = p$. I then use the requirement of common knowledge of rationality to find the equilibrium beliefs, and equilibrium actions of both players, such that these beliefs are consistent given the partitions defined in expressions (34) and (35).

The RCE studied in Chapter 3 is consistent with this definition given the following voter's beliefs

$$\lambda_V(t_V) = 1 \tag{40}$$

Where $\xi(t_V) = \omega_V = (\tilde{u}, (\tilde{a}_1, \tilde{a}_2), G_V, p)$, for some $\tilde{u} \in \bar{U}$ and $(\tilde{a}_1, \tilde{a}_2) \in A$, $\lambda_V(t_V)[R_I^{\mathcal{B}}] = 1$, $\lambda_V(t_V)[C_I^{\mathcal{B}}] = 1$, $\lambda_V(t_V)[\lambda_I(t_V) = 1] = 1$.[9]

[9]The last condition restricts the second-order beliefs of the voter to be degenerate over the event that the incumbent has the same first-order beliefs as the voter. In other words, I restrict the voter to believe that the incumbent has the same beliefs as the voter over the equilibrium strategies, and the fundamentals of the game.

Bibliography

Achen, C. & Bartels, L. (2004), 'Why shark attacks are bad for democracy.', *Unpublished manuscript* .

Alonso, R. & Câmara, O. (2016), 'Bayesian persuasion with heterogeneous priors', *Journal of Economic Theory* **165**, 672 – 706.

Alonso, R. & Câmara, O. (2018), 'On the value of persuasion by experts', *Journal of Economic Theory* **174**, 103–23.

Alpert, M. & Raiffa, H. (1982), A progress report on the training of probability assessors, *in* D. Kahneman, P. Slovic & A. Tversky, eds, 'Judgment Under Uncertainty: Heuristics and Biases', Cambridge University Press.

Alt, J., Mesquita, E. B. d. & Rose, S. (2011), 'Disentangling accountability and competence in elections: Evidence from u.s. term limits', *The Journal of Politics* **73**(1), pp. 171–186.

Argenziano, R. & Weeds, H. (2017), 'Bias and negligence with freedom of information', *Unpublished manuscript* .

Ashworth, S. (2005), 'Reputational dynamics and political careers', *The Journal of Law, Economics, & Organization* **21**(2), 441–466.

Ashworth, S. (2012), 'Electoral accountability: Recent theoretical and empirical work', *Annual Review of Political Science* **15**(1), 183–201.

Ashworth, S. & Bueno De Mesquita, E. (2014), 'Is voter competence good for voters?: Information, rationality, and democratic performance', *American Political Science Review* **108**, 565–587.

Ashworth, S., Bueno de Mesquita, E. & Friedenberg, A. (2018), 'Learning about voter rationality', *American Journal of Political Science* **62**(1), 37–54.

Au, P. H. (2015), 'Dynamic information disclosure', *The RAND Journal of Economics* **46**(4), 791–823.

Austen-Smith, D. (1993), 'Information and influence: Lobbying for agendas and votes', *American Journal of Political Science* **37**(3), 799–833.

Austen-Smith, D. (1998), 'Allocating access for information and contributions', *The Journal of Law, Economics, & Organization* **14**(2), 277–303.

Austen-Smith, D. & Banks, J. (1989), Electoral accountability and incumbency, *in* P. Ordeshook, ed., 'Models of Strategic Choice in Politics', University of Michigan Press.

Austen-Smith, D. & Wright, J. R. (1992), 'Competitive lobbying for a legislator's vote', *Social Choice and Welfare* **9**(3), 229–57.

Barro, R. (1973), 'The control of politicians: An economic model', *Public Choice* **14**(1), 19–42.

Bartels, L. (1996), 'Uninformed votes: information effects in presidential elections.', *American Journal of Political Science* **40**(1).

Basu, P. (2017), 'Dynamic bayesian persuasion with a privately informed receiver', *Unpublished manuscript* .

Bausch, A. & Zeitzoff, T. (2014), 'Citizen information, electoral incentives, and provision of counter-terrorism: An experimental approach', *Political Behavior* pp. 1–26.

Bendor, J., Kumar, S. & Siegel, D. A. (2010), 'Adaptively rational retrospective voting', *Journal of Theoretical Politics* **22**(1), 26–63.

Benesch, C., Butler, M. & Hofer, K. E. (2018), 'Transparency in parliamentary voting', *Journal of Public Economics* **163**.

Bennedsen, M. & Feldmann, S. E. (2002a), 'Lobbying and legislative organization: The effect of the vote of confidence procedure', *Business and Politics* **4**(2), 187203.

Bennedsen, M. & Feldmann, S. E. (2002b), 'Lobbying legislatures', *Journal of Political Economy* **110**(4), 919–46.

Besley, T. (2004), 'Joseph schumpeter lecture: Paying politicians: Theory and evidence', *Journal of the European Economic Association* **2**(2/3), pp. 193–215.

Besley, T. (2007), *Principled agents?: The political economy of good government*, Oxford University Press.

Besley, T. & Burgess, R. (2002), 'The political economy of government responsiveness: Theory and evidence from India', *The Quarterly Journal of Economics* **117**(4), 1415–1451.

Besley, T. & Case, A. (1995*a*), 'Does Electoral Accountability Affect Economic Policy Choices? Evidence from Gubernatorial Term Limits', *The Quarterly Journal of Economics* **110**(3), 769–98.

Besley, T. & Case, A. (1995*b*), 'Incumbent behavior: Vote-seeking, tax-setting, and yardstick competition', *The American Economic Review* **85**(1), pp. 25–45.

Besley, T. & Case, A. (2003), 'Political institutions and policy choices: Evidence from the united states', *Journal of Economic Literature* **41**(1), 7–73.

Bhandari, A., Larreguy, H. & Marshall, J. (2018), 'An empirical anatomy of political accountability: Experimental evidence from a pre-election information dissemination campaign in senegal', *Unpublished manuscript* .

Bhattacharya, S. & Mukherjee, A. (2013), 'Strategic information revelation when experts compete to influence', *The RAND Journal of Economics* **44**(3), 522–544.

Bidner, C. & Francois, P. (2013), 'The emergence of political accountability*', *The Quarterly Journal of Economics* **128**(3), 1397–1448.

Bimber, B. (1991), 'Information as a factor in congressional politics', *Legislative Studies Quarterly* **16**(4), 585–605.

Block, R. A. & Harper, D. R. (1991), 'Overconfidence in estimation: Testing the anchoring-and-adjustment hypothesis', *Organizational Behavior and Human Decision Processes* **49**(2), 188 – 207.

Boleslavsky, R. & Cotton, C. (2018), 'Limited capacity in project selection: competition through evidence production', *Economic Theory* **65**(2), 385–421.

Bonfiglioli, A. & Gancia, G. (2013), 'Uncertainty, electoral incentives and political myopia', *The Economic Journal* **123**(568), 373–400.

Campbell, A., Converse, P., Miller, W. & Stokes, D. (1960), *The American Voter.*, Wiley & Sons.

Caplan, B., Crampton, E., Grove, W. A. & Somin, I. (2013), 'Systematically biased beliefs about political influence: Evidence from the perceptions of political influence on policy outcomes survey', *PS: Political Science & Politics* **46**, 760–767.

Carey, J. M. (2013), 'Transparency & legislative behavior', *Unpublished manuscript* .

Cho, I.-K. & Kreps, D. M. (1987), 'Signaling games and stable equilibria', *The Quarterly Journal of Economics* **102**(2), 179–221.

Clark, C. (2016), 'Should Congressional Research Service Reports Be Kept Secret? Government Executive 20/06/2016, retrieved on 06/08/2018 from: https://www.govexec.com/oversight/2016/06/should-congressional-research-service-reports-be-kept-secret/129227/', *Government Executive* .

Cotton, C. S. & Dellis, A. (2016), 'Informational lobbying and agenda distortion', *The Journal of Law, Economics, & Organization* **32**(4), 762–93.

Cotton, C. S. & Li, C. (2018), 'Clueless politicians: On policymaker incentives for information acquisition in a model of lobbying', *The Journal of Law, Economics, & Organization* **34**(3), 425–56.

Dahm, M. & Porteiro, N. (2008a), 'Informational lobbying under the shadow of political pressure', *Social Choice and Welfare* **30**(4), 531–59.

de Figueiredo, J. M. & Silverman, B. S. (2006), 'Academic earmarks and the returns to lobbying', *The Journal of Law and Economics* **49**(2), 597–625.

DeBonis, M. (2017), 'House panel poised to order that congressional research reports be made public. The Washington Post 28/06/2017, retrieved on 29/03/2018 from: https://www.washingtonpost.com/news/powerpost/wp/2017/06/28/house-panel-poised-to-order-that-congressional-research-reports-be-made-public/?utm_term=.53e4f39dade8', *Washington Post* .

Delli Carpini, M. & Keeter, S. (1996), *What Americans Know about Politics and Why It Matters.*, Yale University Press.

Dellis, A. & Oak, M. (2018), 'Subpoena power and informational lobbying', *Unpublished manuscript* .

Ellis, C. J. & Groll, T. (2017), 'Strategic legislative subsidies: Informational lobbying and the cost of policy', *Unpublished manuscript* .

Esponda, I. (2013), 'Rationalizable conjectural equilibrium: A framework for robust predictions', *Theoretical Economics* **8**(2), 467–501.

Esterling, K. M. (2004), *The political economy of expertise: Information and efficiency in American national politics*, University of Michigan Press.

Fearon, J. D. (1999), Electoral accountability and the control of politicians: Selecting good types versus sanctioning poor performance, *in* A. Przeworski, S. C. Stokes & B. Manin, eds, 'Democracy, Accountability, and Representation', Cambridge University Press, pp. 55–97. Cambridge Books Online.

Fehrler, S. & Hughes, N. (2018), 'How transparency kills information aggregation: theory and experiment', *American Economic Journal: Microeconomics* **10**(1), 181–209.

Felgenhauer, M. (2010), 'Transparency and special interests', *Economics Letters* **106**(1), 4–7.

Felgenhauer, M. (2013), 'Informational and monetary lobbying: Expert politicians, good decisions?', *Journal of Public Economic Theory* **15**(1), 125–55.

Ferejohn, J. (1986), 'Incumbent performance and electoral control', *Public Choice* **50**(1-3), 5–25.

Ferraz, C. & Finan, F. (2011), 'Electoral accountability and corruption: Evidence from the audits of local governments', *American Economic Review* **101**(4), 1274–1311.

Fox, J. (2007), 'Government transparency and policymaking', *Public Choice* **131**(1), 23–44.

Fox, J. & Van Weelden, R. (2012), 'Costly transparency', *Journal of Public Economics* **96**(1), 142–50.

Gailmard, S. & Patty, J. W. (2018), 'Giving advice versus making decisions: Transparency, information, and delegation', *Political Science Research and Methods* .

Ganuza, J.-J. & Penalva, J. S. (2010), 'Signal orderings based on dispersion and the supply of private information in auctions', *Econometrica* **78**(3), 1007–1030.

Gentzkow, M. & Kamenica, E. (2014), 'Costly persuasion', *American Economic Review* **104**(5), 457–62.

Gentzkow, M. & Kamenica, E. (2017*a*), 'Bayesian persuasion with multiple senders and rich signal spaces', *Games and Economic Behavior* **104**, 411 – 429.

Gentzkow, M. & Kamenica, E. (2017*b*), 'Competition in persuasion', *Review of economic studies* **84**(1), 300–322.

Gilligan, T. W. & Krehbiel, K. (1987), 'Collective decisionmaking and standing committees: An informational rationale for restrictive amendment procedures', *The Journal of Law, Economics, & Organization* **3**(2), 287–335.

Gilligan, T. W. & Krehbiel, K. (1989), 'Asymmetric information and legislative rules with a heterogeneous committee', *American Journal of Political Science* **33**(2), 459–90.

Gilligan, T. W. & Krehbiel, K. (1990), 'Organization of informative committees by a rational legislature', *American Journal of Political Science* **34**(2), 531–64.

Gradwohl, R. & Feddersen, T. (2018), 'Persuasion and transparency', *The Journal of Politics* **80**(3), 903–15.

Grossman, S. J. (1981), 'The informational role of warranties and private disclosure about product quality', *The Journal of Law and Economics* **24**(3), 461–83.

Guo, Y. & Shmaya, E. (2018), 'The interval structure of optimal disclosure', *Unpublished manuscript* .

Hansen, S., McMahon, M. & Prat, A. (2017), 'Transparency and deliberation within the fomc: a computational linguistics approach', *The Quarterly Journal of Economics* **133**(2), 80170.

Harrington, J. E. J. (1993), 'The impact of reelection pressures on the fulfillment of campaign promises', *Games and Economic Behavior* **5**(1), 71 – 97.

HC Deb (01 Nov 2017), 'Vol. 630, Col. 926, available at `https://hansard.parliament.uk/commons/2017-11-01/debates/62139AE8-AA75-42F7-85D3-BB50087ACEFB/ExitingTheEUSectoralImpactAssessments` (Accessed: 08 August 2018)', *Hansard* .

Healy, A., Malhotra, N. & Mo, C. (2010), 'Irrelevant events affect voters' evaluations of government performance', *Proceedings of the National Academy of Science USA* **107**(29).

Hedlund, J. (2017), 'Bayesian persuasion by a privately informed sender', *Journal of Economic Theory* **167**, 229 – 268.

Hobolt, S., Tilley, J. & Wittrock, J. (2013), 'Listening to the government: How information shapes responsibility attributions', *Political Behavior* **35**(1), 153–174.

Howlett, M. (2015), 'Policy analytical capacity: The supply and demand for policy analysis in government', *Policy and Society* **34**(3-4), 173–82.

Huber, G. A., Hill, S. J. & Lenz, G. S. (2012), 'Sources of bias in retrospective decision making: Experimental evidence on voters limitations in controlling incumbents', *American Political Science Review* **106**, 720–741.

Igan, D. & Mishra, P. (2014), 'Wall Street, Capitol Hill, and K Street: Political influence and financial regulation', *The Journal of Law and Economics* **57**(4), 1063–84.

Islam, R. (2006), 'Does more transparency go along with better governance?', *Economics & Politics* **18**(2), 121–67.

Jemmer, H. (2014), Lobbying in estonia, mapping the players, risks and political context, Technical report, Transparency International Estonia.

Kamenica, E. & Gentzkow, M. (2011), 'Bayesian persuasion', *American Economic Review* **101**(6), 2590–2615.

Kang, K. (2016), 'Policy influence and private returns from lobbying in the energy sector', *The Review of Economic Studies* **83**(1), 269–305.

Kappe, R. (2013), Asymmetric evaluations: Government popularity and economic performance in the united kingdom, *in* 'EPOP conference, Lancaster University, UK'.

Kartik, N., Lee, F. X. & Suen, W. (2017), 'Investment in concealable information by biased experts', *The RAND Journal of Economics* **48**(1), 24–43.

Kearns, C. E., Schmidt, L. A. & Glantz, S. A. (2016), 'Sugar industry and coronary heart disease research: a historical analysis of internal industry documents', *JAMA internal medicine* **176**(11), 1680–85.

Koehler, D. J. & Harvey, N. (2008), *Blackwell handbook of judgment and decision making*, John Wiley & Sons.

Kolotilin, A. (2018), 'Optimal information disclosure: A linear programming approach', *Theoretical Economics* **13**(2), 607–35.

Kolotilin, A., Mylovanov, T., Zapechelnyuk, A. & Li, M. (2017), 'Persuasion of a privately informed receiver', *Econometrica* **85**(6), 1949–64.

Kopp, E., Lupkin, S. & Lucas, E. (2018), 'Patient Advocacy Groups Take In Millions From Drugmakers. Is There A Payback?. 06/04/2018, retrieved on 28/10/2018 from: `https://khn.org/news/patient-advocacy-groups-take-in-millions-from-drugmakers-is-there-a-payback/`', *Kaiser Health News* .

Krehbiel, K. (1991), *Information and Legislative Organization*, Michigan Studies In Political Analysis, University of Michigan Press.

Krehbiel, K. (2004), 'Legislative organization', *Journal of Economic Perspectives* **18**(1), 113–28.

Krishna, V. & Morgan, J. (2001), 'A model of expertise', *The Quarterly Journal of Economics* **116**(2), 747–775.

Lagerlof, J. (1997), 'Lobbying, information, and private and social welfare', *European Journal of Political Economy* **13**(3), 615–37.

Lau, R. R. & Redlawsk, D. P. (2001), 'Advantages and disadvantages of cognitive heuristics in political decision making', *American Journal of Political Science* **45**(4), pp. 951–971.

Leigh, A. (2009), 'Does the world economy swing national elections?', *Oxford Bulletin of Economics and Statistics* **71**(2).

Levy, G. (2007), 'Decision making in committees: Transparency, reputation, and voting rules', *American Economic Review* **97**(1), 150–68.

Levy, G., Moreno De Barreda, I. & Razin, R. (2018), 'Media power via the correlation of news content', *Unpublished manuscript* .

Levy, G. & Razin, R. (2014), 'Does polarization of opinions lead to polarization of platforms? the case of correlation neglect', *Unpublished manuscript* .

Levy, G. & Razin, R. (2015), 'Correlation neglect, voting behaviour and information aggregation', *American Economic Review* **105**(4), 1634–45.

Li, F. & Norman, P. (2018), 'On bayesian persuasion with multiple senders', *Economics Letters* **170**, 66 – 70.

Li, H. & Shi, X. (2017), 'Discriminatory information disclosure', *American Economic Review* **107**(11), 3363–85.

Lockwood, B. (2017), 'Confirmation bias and electoral accountability', *Quarterly Journal of Political Science* **11**(4), 471–501.

Malesky, E., Schuler, P. & Tran, A. (2012), 'The adverse effects of sunshine: A field experiment on legislative transparency in an authoritarian assembly', *American Political Science Review* **106**(4), 762–86.

Martimort, D. & Semenov, A. (2008), 'The informational effects of competition and collusion in legislative politics', *Journal of Public Economics* **92**(7), 1541–1563.

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

Matthews Luxon, E. O. (2012), Lobbying in the dark?: the effects of policy-making transparency on interest group lobbying strategies in France and Sweden, PhD thesis, UC San Diego.

Meade, E. & Stasavage, D. (2008), 'Publicity of Debate and the Incentive to Dissent: Evidence from the US Federal Reserve', *The Economic Journal* **118**(528), 695–717.

Meirowtiz, A. (2007), 'Probabilistic voting and accountability in elections with uncertain policy constraints', *Journal of Public Economic Theory* **9**(1), 41–68.

Milgrom, P. & Roberts, J. (1986), 'Relying on the information of interested parties', *The RAND Journal of Economics* **17**(1), 18–32.

Moore, D. A. & Healy, P. J. (2008), 'The trouble with overconfidence.', *Psychological review* **115**(2), 502.

Naurin, D. (2007), *Deliberation behind closed doors: transparency and lobbying in the European Union*, ECPR press.

Nestle, M. (2016), 'Food industry funding of nutrition research: the relevance of history for current debates', *JAMA internal medicine* **176**(11), 1685–86.

Ortoleva, P. & Snowberg, E. (2015), 'Overconfidence in political behavior', *American Economic Review* **105**(2), 504–35.

Padro i Miquel, G. & Snowberg, E. (2012), 'The lesser evil: Executive accountability with partisan supporters', **24**(1), 19–45.

Payson, J. A. (2017), 'Cities in the statehouse: How local governments use lobbyists to secure state funding', *Unpublished manuscript* .

Perez-Richet, E. (2014), 'Interim bayesian persuasion: First steps', *American Economic Review* **104**(5), 469–74.

Persson, T., Roland, G. & Tabellini, G. (1997), 'Separation of powers and political accountability', *The Quarterly Journal of Economics* **112**(4), 1163–1202.

Potters, J. & van Winden, F. (1992), 'Lobbying and asymmetric information', *Public Choice* **74**(3), 269–92.

Prat, A. (2005), 'The wrong kind of transparency', *The American Economic Review* **95**(3), 862–77.

Rasmusen, E. (1993), 'Lobbying when the decisionmaker can acquire independent information', *Public Choice* **77**(4), 899–913.

Richter, B. K., Samphantharak, K. & Timmons, J. F. (2009), 'Lobbying and taxes', *American Journal of Political Science* **53**(4), 893–909.

Schnakenberg, K. E. (2017), 'Informational lobbying and legislative voting', *American Journal of Political Science* **61**(1), 129–45.

Seidmann, D. J. (2011), 'A theory of voting patterns and performance in private and public committees', *Social Choice and Welfare* **36**(1), 49–74.

Shin, H. S. (1998), 'Adversarial and inquisitorial procedures in arbitration', *The RAND Journal of Economics* **29**(2), 378–405.

Snyder, J. M. J. & Stromberg, D. (2010), 'Press coverage and political accountability', *Journal of Political Economy* **118**(2), pp. 355–408.

Snyder, J. M. & Ting, M. M. (2008), 'Interest groups and the electoral control of politicians', *Journal of Public Economics* **92**(3), 482–500.

Stasavage, D. (2004), 'Open-door or closed-door? transparency in domestic and international bargaining', *International Organization* **58**(4), 667–703.

Stasavage, D. (2007), 'Polarization and publicity: Rethinking the benefits of deliberative democracy', *Journal of Politics* **69**(1), 59–72.

Stephenson, M. C. (2011), 'Information acquisition and institutional design', *Harv. L. Rev.* **124**, 1422–83.

Swank, O. H. & Visser, B. (2013), 'Is transparency to no avail?', *The Scandinavian Journal of Economics* **115**(4), 967–94.

White, J. & Bero, L. A. (2010), 'Corporate manipulation of research: strategies are similar across five industries', *Stanford Law & Policy Review* **21**, 105–33.

Wolfers, J. (2007), 'Are voters rational?: Evidence from gubernatorial elections', *Unpublished manuscript* .

Wolton, S. (2018), 'Lobbying, inside and out: How special interest groups influence policy choices', *Unpublished manuscript* .