On the (Mis-)Use of Intelligence for Public Debate

Andrea Patacconi

University of Oxford

April 2006

Abstract: Many observers have expressed concern that, in the case of Iraq, intelligence assessments may have been warped by inappropriate political pressure. Survey evidence also indicates that many believe that official documents and statistics are subject to political interference. This paper examines these issues in a model where (i) the preferences of the government and the public are not perfectly aligned, (ii) the government wants to be perceived as taking the 'right' decision by the public, and (iii) the process of information gathering and evaluation can be manipulated. The analysis sheds some light upon the issue of transparency in government. In a disclosure environment, policymakers face a tradeoff between preserving the quality of the information brought to bear on a decision and the benefits of distorting the information made available for public debate. In a nondisclosure environment, by contrast, the government has no incentive to manipulate the information. Paradoxically, a policy of disclosure can sometimes benefit the public exactly because it creates incentives for manipulation.

Keywords: Transparency in Government, Disclosure, Garbling of Information, Political Accountability, Intelligence.

JEL Classification: D73, H11, H56.

[1]Postal address: Andrea Patacconi, Department of Economics, Manor Road Building, Manor Road, Oxford, OX1 3UQ, United Kingdom. E-mail: andrea.patacconi@economics.ox.ac.uk.

I would like to thank Karen Croxson, Diego Gambetta, Ian Jewitt, Kohei Kawamura, seminar participants at the University of Oxford, and especially Clare Leaver and Meg Meyer for helpful discussions and remarks. Financial support from the Economics Department at the University of Oxford and the Royal Economic Society is gratefully acknowledged. All remaining errors are mine.

The assessment process [...] must avoid being so captured by policy objectives that it reports the world as policy-makers would wish it to be rather than it is" (Butler Report, p.16)

We also recognize that there is a real dilemma between giving the public an authoritative account of the intelligence picture and protecting the objectivity of the JIC [Joint Intelligence Committee] from the pressures imposed by providing information for public debate. (Butler Report, p.114)

1 Motivation

The last few years have witnessed an explosion of interest in the workings of the intelligence community. One of the main reasons for this has been the tremendous hit taken by America and Britain's intelligence services when Iraq's expected stock of weapons of mass destruction (WMD) failed to materialize. Prompted by that failure, the British and the American administrations established three commissions to review the gathering, evaluation and use of intelligence on WMD, and to make recommendations for the future.[1]

These commissions have highlighted a number of possible explanations for the errors contained in the U.K. and U.S. Intelligence Community's Prewar Intelligence Assessment on Iraq. However, they all put the greatest emphasis on the risks generated by the intelligence community's collective presumption that Iraq had an active and growing WMD program. According to the Senate Intelligence Committee, for instance, it was this presumption, and the resulting 'group think' dynamic, that led intelligence community analysts, collectors and managers both to interpret ambiguous evidence as conclusively indicative of a WMD program as well as to ignore or minimize evidence that Iraq did not have active and expanding weapons of mass destruction programs (Conclusion 3).[2]

In this paper, we set aside psychological factors such as 'group think' and focus instead

[1] Their results are contained in three reports, the Review of Intelligence on Weapons of Mass Destruction (2004) for the U.K. (henceforth, the Butler Report) and, for the U.S., the Report of the Select Committee on Intelligence on the U.S. Intelligence Community's Prewar Intelligence Assessments on Iraq (2004) and the Report of the Commission on the Intelligence Capabilities of the United States Regarding Weapons of Mass Destruction (2005). In the following, we will refer to the last two documents as the reports of the Senate Intelligence Committee and of the WMD Commission, respectively.

[2] See Garicano and Posner (2005) for an excelent discussion of the organizational issues involved in intelligence collection and analysis.

on the risk of capture of the intelligence community. Many observers have in fact expressed concern that, in the case of Iraq, intelligence community judgements may have been warped by inappropriate political pressure, so that intelligence could be used to build a case for war. The role of intelligence in shaping public opinion was clearly understood by politicians and the media. There is also evidence that some senior intelligence analysts knew what their political masters wanted and were ready to provide it. Especially revealing is the case of Curve Ball, a drunkard whose claims constituted the backbone of the intelligence on Iraq's mobile biological weapons program mentioned by the U.S. Secretary of State Colin Powell in February 2003 at the U.N. Security Council. When, before Mr. Powell delivered his speech, the CIA agent who had interviewed Curve Ball raised concerns about his reliability, he was told by the Deputy Chief of the CIA's Iraqi Task Force:

> "As I said last night, let's keep in mind the fact that this war's going to happen regardless of what Curve Ball said or didn't say, and that the Powers That Be probably aren't terribly interested in whether Curve Ball knows what he's talking about."[3]

More generally, a major concern is that analysts may engage in unprofessional manipulation of information and judgements to please their superiors if they believe that such behavior will be rewarded. In that respect, it is not encouraging that "the analysts who raised concerns about the need for reassessments were not rewarded for having done so but were instead forced to leave WINPAC" (WMD Commission, p.193).[4]

Motivated by these considerations, the present paper shows that the 'politicization' of intelligence may be instrumental in winning support for the policy that is eventually chosen by the government when the preferences of the policymakers and those of the public are not fully aligned. The drawback we emphasize is that the quality of the information brought to bear on the decision might be compromised. As Krueger and Laitin (2004) put it, in fact, "only accurate information, presented without political spin, can help the public and decision-makers know where the United States stands in the war on terrorism and how best to fight it" (p.13). This paper then examines whether much pre-decision information should be kept secret, and highlights some of the costs

[3]Extract from an e-mail provided to the WMD Commission, p.249.

[4]For more evidence consistent with the idea of politicization of intelligence, see The Economist's 'Special Report Intelligence Failures' (17/07/04, pp.23-25). However, it should be noted that the British and the American commissions concluded that no analytic judgement had been made or changed by the intelligence community in response to political pressure to reach a particular conclusion.

(in terms of manipulation of information) and benefits (in terms of greater accountability) of transparency in government.

Before proceeding, however, it should be pointed out that the scope of the present analysis is not limited to the problem of intelligence failures, but applies more generally to situations where disclosure of information creates incentives for manipulation. In politics, for instance, there is a widespread belief that 'transparency' instruments such as the Freedom of Information Act might have negative effects on frank conversation inside the government. Roberts (2006), in particular, discusses several bureaucratic practices, including changes in recordkeeping, decline in candor, manipulation of records, failure to create records, that have had the effect of undercutting the right to information. He reports that, to limit the damage done by the release of information under the law and acting on instructions from the Prime Minister's office, the Canadian Department of Public Works kept "minimum information" on the Sponsorship Program spending and later created expenditure guidelines for cosmetic purposes. Roberts' findings clearly suggest that transparency may lead to biased information gathering.

Another point worth mentioning is that one of the most common exemptions to the principle of open government in practice is the exemption of pre-decision information. According to Frankel (2001), in fact, "Member states all to a greater or lesser degree protect the opinions, advice or the exchange of views which take place during policy-making" (p.8). Such protection is typically motivated by the concern that disclosure of information may temper the frankness and candor of discussion inside the government to the detriment of the decision making process. The present model provides support for the idea that secrecy is effective in preserving the integrity of the decision making process. However, it also shows that secrecy is rather ineffective in keeping policymakers accountable and thus responsive to public desires.[5]

Finally, the present paper might shed some light upon the controversial use of official statistics in Britain. As The Economist put it: "The more that ministerial reputations ride on statistics, the more protection statisticians need from political interference. Current arrangements fail to

[5] The degree to which pre-decision material is protected under the law varies considerably across countries. At one extreme "The Danish law has a wide class exemption for internal case material, including documents prepared by the authority for its own use. These remain exempt after decisions are taken" (Frankel (2001), p.8). In Sweden, instead, pre-decision information must be disclosed after a decision has been "finally settled". In the present context, what matters is that delays can substantially reduce the value of released information, in particular its potentially damaging impact for policymakers.

deliver that protection." (26/03/05, p.31-32).[6] In that context, this paper points to a potential tradeoff between biasing the statistics to show the public that government policies are working and preserving their integrity, so that they can be used to inform policymakers.

2 Overview

In this section we present an overview of the model and discuss its key assumptions.

The model is a very stylized one. There are two players, a government (policymakers, agent, etc.) and the public (citizens, principal, etc.). The government must choose whether or not to implement a policy which affects the welfare of both parties. The outcome of this choice depends on an unknown state of the world, but the government also enjoys a private benefit when the policy is implemented. This creates a potential conflict of interest between the government and the public. Before taking its decision, the government collects some information about the state of the world.

Citizens play no explicit role in the decision making process. Nevertheless, their opinion matters because the government wants to be perceived as taking the 'right' decision by the public, say because it wants to improve its political standing and/or electoral chances. In this way we capture the disciplining effect of public opinion on policy and measure the government's electoral (or legitimacy) concerns.

The public bases its decision to support implementation or the status quo on the information made available for public debate. However, before information is collected, the government can distort the process of information gathering and evaluation, should that be in its interest. The specific idea we explore is that the government may allow fake signals that support its initial disposition (relative to the public) percolate through the system. Examples might be the government seeking the advice of 'experts' who are known to be biased in favor of a particular policy, biased information evaluation or asymmetric vetting. In line with the last two interpretations, we will sometimes say that the government can bias the 'screening technology', and will refer to the

[6] In line with these concerns, a recent survey by the Office for National Statistics in Britain shows that the majority of respondents believed official figures were changed to support a particular argument (68%), that there was political interference in their production (58%), and that mistakes were suppressed (69%) (Jones and Jones (2005)). A survey by the British Medical Association (2005) also found that only 26% of accident and emergency (A&E) departments thought the figures they submitted for the waiting-time target were a true reflection of their performance, and 16% reported direct manipulation of data.

extent to which information is manipulated as the 'size of the bias'.[7] The drawback we emphasize is that, in an attempt to justify a certain course of action, the quality of the information brought to bear on a decision may be compromised.

Information can be disclosed to the public. The key assumption we make is that, once potentially biased information is gathered, it must be reported truthfully, if reported at all. One might think of a situation where policymakers can be severely punished if caught lying, but blame can be placed on bureaucrats (or informants) should the information turn out to be forged.[8] We also rule out 'presentational strategies' such as spin and exaggerations, since they may not be particularly effective when citizens are sophisticated (as in this model) and the press is free and diverse (Mullainathan and Shleifer (2004)).

This paper studies a number of disclosure environments. At one extreme, the government may be forced to disclose all the information (full disclosure). Even in this case, however, an important informational problem arises, depending on whether the size of the bias in information (optimally chosen by the government) is observed by the public or only inferred in equilibrium. After introducing the model in Section 3, we begin with the case where the size of the bias is observed by the public (Section 4). This could be because not only the reports but also the identity of the experts (and hence their reputations) are disclosed, or because the press provides enough information–for instance allegations of political interference–that the public is able to form a precise opinion about the level of politicization in government agencies. Our main result is that policymakers then face a tradeoff between preserving the quality of the information brought to bear on the decision and the benefits of distorting the information made available for public debate. Interestingly, we also show that, as the quality of information deteriorates, the government's chosen policy tends to more closely reflect citizens' preferences (discipline). In fact, as information becomes more biased toward the government's initial disposition, policymakers become increasing concerned about forgeries and therefore tend to implement the project less

[7] That policymakers can easily interfere with the process of information gathering and evaluation is obviously a simplification. Nevertheless, it has long been recognized that "Senior civil servants are increasingly called on to expand their minister's policies publicly [...] What is more serious is the use made of senior civil servants to sell government policies, thus the government's political standing and (it hopes) its electoral chances." (Ridley (1991) p.447). In the US, for instance, "Congress can use appropriations as both carrot and stick—providing additional funds for bureaucrats who produce pleasing decisions and withholding funds for those who do not" (Arnold (1987), p.280). We briefly comment on the important issue of the independence of the bureaucracy in the concluding remarks.

[8] In this model, the public does not penalize the government just for distorting the screening technology. This is a strong assumption, but one that could be easily relaxed. See the concluding remarks.

often, which is what the public wants. Thus, manipulations turn out to have a bright side, for they help more closely align the government's interests with those of the public.

Section 5 studies the case where the size of the bias cannot be directly observed by the public. It shows that the analysis of Section 4 is robust provided that the conflict of interest between the public and the government is not 'too large'. When the conflict of interest is very large, no pure strategy Nash equilibria exist.

Section 6 considers a nondisclosure scenario (termed 'secrecy') where the government commits not to disclose any information and the size of the bias in information is unobservable. Nevertheless, the public can update its beliefs by observing the government's choice. In line with conventional wisdom, we show that the government has no incentive to manipulate the information. However, since the public cannot in general figure out exactly why a specific decision was taken, accountability is compromised and the government becomes less responsive to public desires, relative to full disclosure. Furthermore, since information is always so reliable, the conflict of interest between the government and the public is accentuated.

Section 7 compares full disclosure to secrecy. It shows that either scenario can be optimal from the public's viewpoint. Transparency performs well when the government is not strongly biased in favor of the project. Secrecy tends to be preferable when disciplining the government is of secondary importance relative to preserving the quality of information. Interestingly, disclosure can be optimal even when disclosure per se (that is, without manipulation of information) is not sufficient to discipline the government. Thus, paradoxically, the public could be worse off if the government committed not to manipulate the information.

Section 8 deals with two extensions. Subsection 8.1 shows that, as in most models with hard information, voluntary disclosure will typically result in all the information being disclosed.[9] Subsection 8.2 studies a variant of the model where decision-making powers are allocated to the public ('direct democracy'). Section 9 concludes. All the proofs are gathered in the Appendix.

2.1 Related Literature

The paper most directly contributes to a growing literature focusing on transparency in principal-agent relationships. The classic reference is Holmstrom (1979). Holmstrom provides a strong

[9] Verrecchia (2001), in particular, shows that in most auditing situations disclosing verifiable information is optimal.

rationale in favor of transparency. In his model, more information about an agent's action makes the agent more accountable and therefore helps more closely align the interests of the agent with those of the principal. Prat (2006) lists several reasons for why transparency may not always be desirable, including the right to privacy, the direct cost of disclosure and the risk that hostile parties learn sensitive information. Exceptions to the Holmstrom's result focusing on purely incentive considerations are provided by Cremer (1995) and Holmstrom (1999), among others. However, the paper that is most closely related to the present one is Prat (2005). Prat explores the idea that a certain kind of transparency may create perverse incentives. He studies a model of career concerns for experts where a principal may observe an agent's action and/or the consequences of that action. His main result is that transparency on action may induce the agent to disregard useful private information and act in a conformist manner.[10] As a result, the principal may be better off by committing not to observe the action. By contrast, transparency on consequences is always beneficial to the principal.

The present paper focuses neither on transparency on action nor on consequences. In this model transparency measures the extent to which pre-decision information is shared between the agent and the principal. Actions are observable but their long-term consequences are not (these are reasonable assumptions in politics and could be relaxed). The key issue is not whether transparency induces conformism on the part of the agent, but whether an agent would distort his own information (and possibly the principal's) to influence how the principal perceives his action.[11] In that respect, the present paper also differs from recent contributions looking at the issue of transparency in group decision making (see Levy (2005) and the references therein). The intuition in these papers is similar to Prat's: Secrecy may induce better decisions because, if individual votes cannot be observed, the members of a committee have less of an incentive to distort their actions in order to signal their types.

Finally, this paper is related to a large literature focusing on accountability in government (see, e.g., Chapter 4 and 9 in Persson and Tabellini (2002) and Maskin and Tirole (2004)). We contribute to this literature by studying how transparency affects both manipulation of information and accountability.

[10] In Prendergast (1993), instead, conformism is generated by subjective performance evaluation, which may induce a worker to bias his report toward the principal's signal.

[11] It should be mentioned that a few papers have looked at the issue of manipulation of information in agency. See, e.g., Maggi and Rodriguez-Clare (1995) and Dewatripont and Tirole (1999).

3 The Model

3.1 Preferences

There are two players, the government (or policymakers) and the public (citizens, voters). The government gathers some information and then decides whether or not to implement a project. The government policy is denoted by $d \in \{w, n\}$, where w stands for "implementation" or "radical action" (e.g., a large increase in the NHS budget, go to war, etc.) and n denotes the "status quo" (an ‘average’ increase in the NHS budget, engage in diplomatic efforts, etc.). Citizens observe d and any disclosed information and then choose which policy $v \in \{w, n\}$ they support. Their choice is made before outcomes are observed.

Players maximize expected payoffs. These payoffs depend on the true state of the world, $S \in \{W, N\} \equiv \Lambda$, which is initially unknown, with both states being a priori equally likely. For instance, W may denote the state of the world in which Iraq has WMD, and N a scenario where Iraq has no such weapons. We assume that both the government and the public incur a loss of C_w when $d = w$ but $S = N$, and a loss of C_n when $d = n$ but $S = W$. Their payoffs when d matches the true state of the world (that is, when the government takes the ‘right’ action) is zero, and with no loss of generality C_w and C_n are normalized so that $C_w + C_n = 1$.

The public’s payoff is determined solely by the above considerations, and therefore does not depend on v, the policy they support. Citizens choose v as follows. Let σ_P denote the public’s belief that $S = W$. (The information structure of the game is discussed later.) Citizens support w if they believe w to be the ‘right’ action, that is, if

$$-C_w (1 - \sigma_P) \geq -C_n \sigma_P \quad \Leftrightarrow \quad \sigma_P \geq C_w \equiv T_P. \tag{1}$$

Similarly, they support n if $\sigma_P < T_P$.[12]

The preferences of the government are more complex. In particular, in addition to the costs and benefits already mentioned, the government also enjoys a benefit $B \geq 0$ when the project is implemented. In the following, B will often be referred to as a private benefit. Note however the present framework might also describe a situation where the government maximizes social welfare, and the public is for some reason biased against the project.[13]

[12] For now, we arbitrarily break ties in favor of w. But see Section 6.

[13] One might also consider a model where there are no private benefits, but players have different prior beliefs.

Finally, to capture the disciplining effect of public opinion, the government is assumed to incur a loss $E \geq 0$ whenever its chosen policy is not supported by the public. Different interpretations can be attached to this cost. The most straightforward one is in terms of electoral losses: if a party adopts an unpopular policy, citizens may vote for an opposition party in the next election. Alternatively, one may consider a situation where a country wants to convince other nations (the "public") that a particular course of action is justified in order to receive logistic or military support. Here E would measure the loss to that country should support be denied. Notice that in these settings v denotes a binary action (e.g., whether or not to vote for the incumbent). Of course, it would also be interesting to study situations where the citizens' action varies proportionally with the intensity of their preferences. That however would complicate the analysis considerably, and is left for future research.

We formalize the above discussion as follows. Let σ_{Gov} denote the government's belief that $S = W$. Let $1_{\{d,v\}}$ be an indicator function taking value 1 if the government policy is not supported by the public ($d \neq v$), and 0 otherwise. The government implements the project if

$$-C_w[1 - \sigma_{Gov}] + B - 1_{\{w,v\}}E \geq -C_n\sigma_{Gov} - 1_{\{n,v\}}E$$

or, equivalently, if

$$\sigma_{Gov} \geq C_w - B + (1_{\{w,v\}} - 1_{\{n,v\}})E. \tag{2}$$

As before, the project is implemented when the probability that W is the true state is sufficiently high. Relative to (1), however, the government's threshold tends to be lower (and the decision more biased toward implementation) because of B. The government will also tend to choose whichever policy happens to be 'popular'.

A particularly important case in the following will be when the government and the public share the same beliefs.

Lemma 1 *Suppose $\sigma_{Gov} = \sigma_P$. When the public supports w, the government will also choose w. However, when the public supports n, the government will choose w if $\sigma_{Gov} \geq T_{Gov} \equiv C_w - B + E$.*

The intuition for this result is straightforward. Since the government is more disposed toward the project than the public is and they share the same information, if there is enough evidence

For simplicity, however, we will stick with the above setup.

to convince the citizens that w is the 'right' action, there is also enough evidence to convince the government. The opposite is of course not necessarily true, because of the private benefit B. Notice that the government is less likely to pick an unpopular policy if E is large. In particular,

Lemma 2 *If the parties share the same information and* $E \geq B$*, the government will always choose the policy that the public supports.*

When the government always chooses the policy that the public supports, we say that the government is disciplined by public opinion.

3.2 Information Structure

Before taking a decision, the government observes two signals, $s_i \in \{\alpha, \varnothing\} \equiv \Omega$, $i = 1, 2$.[14] An α-signal provides evidence in support of W and suggests that the project should be implemented. A $\varnothing$-signal provides evidence favorable to N and the status quo. The process that generates these signals can be distorted by the government. The specific idea we explore is that the government may let fake α-signals percolate through the system. That could be in the government's interest since policymakers are biased toward the project, relative to the public, and captures in a simple and intuitive way the idea of asymmetric vetting.

This idea is formalized as follows. We begin with two genuine signals, denoted by s_i^G (or s_i^0). Genuine signals are informative and conditionally independent. In particular, $\Pr(s_i^G = \alpha|W) = \Pr(s_i^G = \varnothing|N) = \theta$, where $\theta \in (\frac{1}{2}, 1)$ measures the precision of the signal. Before these signals are observed, $\varnothing$-signals can be transformed into fake α-signals. The probability that a genuine $\varnothing$-signal is transformed into a fake α-signal is $q \in [0, 1]$. Formally, $\Pr(s_i^q = \alpha \mid s_i^G = \varnothing) = q$, where $\mathbf{s}^q = \{s_1^q, s_2^q\}$ denote the observed signals. (Forgeries are independent so that the conditional probability of two signals being forged is q^2.) The probability that a genuine α-signal is transformed into a fake $\varnothing$-signal is zero.[15] Thus, q measures how biased the process of information gathering and evaluation is in favor of the government's initial disposition (relative to the public). The 'size of the bias' q is chosen by the government to maximize its expected payoff,

[14] In this model two signals (or one signal with three possible realizations) are needed because we want to capture a situation where a government is willing to take radical steps on the basis of mixed evidence, but the public only supports those steps when the evidence is clear-cut. Notice that since the number of signals is given, we abstract from the important issue of how much effort should be devoted to collect information.

[15] The analysis of a more general model where fake $\varnothing$-signals can also percolate through the system is conceptually straightforward but analytically cumbersome, and is therefore left for future research.

with all q's being equally costly. The latter assumption allows us to focus on the incentives to manipulate the information for purely electoral reasons.

This simple information structure can be interpreted in many ways. An example is a government seeking the advice of an expert who is known to be biased in favor of the project. The expert may succeed in forging an α-signal when he actually got a $\varnothing$-signal with probability q. Thus q can be interpreted as a measure of how biased the expert is. A second example is asymmetric vetting. Suppose again that information can be forged. Now, however, suppose that an agency must screen the signals to weed out forged information. In such setting, q would measures how carelessly the agency questions information that supports the government's initial disposition (forged $\varnothing$-signals are always detected and disregarded).

Before proceeding, we stress that the government only observes $\mathbf{s}^q$, not $\mathbf{s}^G$, unless it picks $q = 0$. Thus, if the government observes $s^q = \alpha$, $q > 0$ and reports α, it might be reporting forged information, but it would not be telling a lie. In this sense we say that disclosure is 'truthful'.

We also emphasize that the notion of manipulation in this paper is an instance of 'garbling' of information as defined by Marschak and Miyasawa (1968). Indeed,

$$\Pr(\mathbf{s}^q, \mathbf{s}^G \mid S) = \Pr(\mathbf{s}^q \mid \mathbf{s}^G) \Pr(\mathbf{s}^G \mid S) \tag{3}$$

for all $\mathbf{s}^G, \mathbf{s}^q \in \Omega^2$ and $S \in \Lambda$. In other words, the conditional probability distribution $\Pr(\mathbf{s}^q \mid \mathbf{s}^G, S)$ is independent of S, capturing the idea that $\mathbf{s}^q$ is determined solely by $\mathbf{s}^G$, but possibly with the intervention of "noise".

4 Full Disclosure Scenario

In this section, the government truthfully discloses all the signals. We begin with the case where the size of the bias q is observed by the public, as in the 'biased reviewers' model by Lerner and Tirole (2006). The case where q cannot be directly observed but is inferred in equilibrium is studied in the next section.

4.1 Efficient vs. Biased Screening Technologies

We first consider the choice between two screening technologies, the 'efficient' technology $q = 0$ and a biased technology $q \in (0, 1)$. The problem of the optimal choice of q is considered later.

Notice that since the government and the public observe the same signals $\mathbf{s}^k \in \Omega^2$, $k = \{0, q\}$, and q is observable, the parties will share the same posterior beliefs. For notational ease, write $(\cdot, \cdot)^k$ for $\mathbf{s}^k = (\cdot, \cdot)$. The following posteriors will be key:

Efficient Screening Technology	Biased Screening Technology
$\sigma_+^G \equiv \Pr(W\|(\alpha, \alpha)^G) = \frac{\theta^2}{V}$	$\sigma_+^q \equiv \Pr(W\|(\alpha, \alpha)^q) = \frac{[\theta+q(1-\theta)]^2}{[\theta+q(1-\theta)]^2+[(1-\theta)+q\theta]^2}$
$\sigma^G \equiv \Pr(W\|(\alpha, \varnothing)^G) = \frac{1}{2}$	$\sigma^q \equiv \Pr(W\|(\alpha, \varnothing)^q) = \frac{[\theta+q(1-\theta)](1-\theta)}{[\theta+q(1-\theta)](1-\theta)+[(1-\theta)+q\theta]\theta}$
$\sigma_-^G \equiv \Pr(W\|(\varnothing, \varnothing)^G) = \frac{(1-\theta)^2}{V}$	$\sigma_-^q \equiv \Pr(W\|(\varnothing, \varnothing)^q) = \frac{(1-\theta)^2}{V}$

where $V \equiv \theta^2 + (1-\theta)^2 = \Pr((\alpha, \alpha)^G) + \Pr((\varnothing, \varnothing)^G)$ and (for future reference) $R \equiv 2\theta(1-\theta) = \Pr((\alpha, \varnothing)^G) + \Pr((\varnothing, \alpha)^G)$.

Clearly, our Bayesian agents are more skeptical about α-signals than of $\varnothing$-signals, since only the former can be forged: $\Pr(W|(\alpha, \varnothing)^k) = \Pr(W|(\varnothing, \alpha)^k) \leq \frac{1}{2}$. Furthermore, α-signals become less and less informative as the screening technology deteriorates: $\frac{\partial \sigma_+^q}{\partial q}, \frac{\partial \sigma^q}{\partial q} < 0$. Notice that $\lim_{q \to 1} \sigma_+^q = \frac{1}{2}$ and $\lim_{q \to 1} \sigma^q = 1 - \theta$. Let $\sigma_-^G = \sigma_-^q = \sigma_-$.

The following assumptions will be maintained throughout the paper:

A1 $T_P \equiv C_w \in (\frac{1}{2}, \sigma_+^G]$

A2 $C_w - B > \sigma_-$

Assumption 1 implies that the public is willing to support w when the evidence in favor of W is strong (i.e., $\mathbf{s}^G = (\alpha, \alpha)$) but not when the evidence is "mixed" (e.g., $(\alpha, \varnothing)^G$). Assumption 2 can be interpreted as a weak form of congruency between the preferences of the government and those of the public. It says that even without the disciplining effect of public opinion, the government will choose the status quo if the evidence points clearly in that direction: $(\varnothing, \varnothing)^G$. No upper bound is imposed on E (and hence on $T_{Gov} = C_w - B + E$).

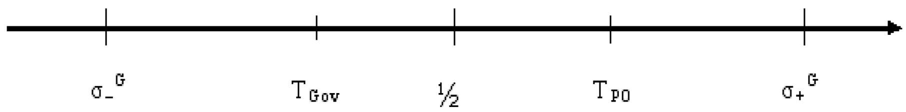

Figure 1: Large Conflict of Interest: $T_{Gov} \leq \frac{1}{2}$

Let $\pi(d, v \mid S)$ denote the payoff to the government when the actions are (d, v) and the state is S. For any given $q \in [0, 1]$, the government's expected payoff is

$$E(\pi^q) = \sum_{\mathbf{s}^q \in \Omega^2} \Pr(\mathbf{s}^q) \underbrace{\sum_{S \in \Lambda} \pi(d(\mathbf{s}^q), v(\mathbf{s}^q) \mid S) \Pr(S \mid \mathbf{s}^q)}_{\equiv \pi(d(\mathbf{s}^q), v(\mathbf{s}^q) \mid \mathbf{s}^q)} \tag{4}$$

where $d(\mathbf{s}^q)$ and $v(\mathbf{s}^q)$ denote, respectively, the choice of the government and that of the public as a function of $\mathbf{s}^q$. Clearly the public can condition its choice on $\mathbf{s}^q$ because the signals are disclosed. This will no longer be true in the nondisclosure scenario.

It is important to note that the government will never choose q so large that the public does not support w when the signals are $(\alpha, \alpha)^q$. Indeed, by selecting $q = 0$, the government would obtain both unbiased information and the backing of the public when $(\alpha, \alpha)^G$. More formally, in the full disclosure scenario (q observable), it must be that in equilibrium $T_P \leq \sigma_+^q$ (see Proposition 3). Let $q^{\max}$ solve

$$\sigma_+^{q^{\max}} = T_P.$$

Then $T_P \leq \sigma_+^q$ is equivalent to $q \in [0, q^{\max}]$. Notice that if $q^{\max}$ tends to 1 as C_w approaches to $\frac{1}{2}$. For ease of exposition, we assume that $q^{\max} = 1$ in the remainder of this subsection.

In the following, it will be useful to distinguish a few cases. The first distinction concerns whether $T_{Gov} \leq \frac{1}{2}$ or $T_{Gov} > \frac{1}{2}$. In the first case, the government implements the project when the (genuine) evidence is mixed, even without the support of the public. Here not even full disclosure is sufficient to discipline the government. Most of the discussion will focus on this case. The $T_{Gov} > \frac{1}{2}$ scenario is studied in Subsection 4.4.

Suppose $T_{Gov} \leq \frac{1}{2}$. Two situations can arise. If q is 'small' ($\sigma^q \geq T_{Gov}$), the government implements the project when the evidence is mixed, e.g., $(\alpha, \varnothing)^q$. We term this scenario Case

A. By contrast, if q is so big that $\sigma^q < T_{Gov}$ (Case B), the government will choose the status quo when the evidence is mixed. Notice a tradeoff here. In Case B, the screening technology is severely biased (q 'large'), but the government is disciplined by public opinion: it changes its decision rule so that its chosen policy is *always* supported by the public. This is not true in Case A. Thus, manipulations may help more closely align the interests of the government with those of the public.

More formally, let $\hat{q}$ solve $\sigma^q = T_{Gov}$ (this requires $\frac{1}{2} \geq T_{Gov} > 1 - \theta$). It can be shown that

$$\hat{q} = r\frac{1/2 - T_{Gov}}{T_{Gov} - \sigma_-}$$

where $r = \frac{R}{V} = \frac{2\theta(1-\theta)}{\theta^2+(1-\theta)^2} < 1$ (see the Appendix). If $T_{Gov} \leq 1 - \theta$, let $\hat{q} = 1$. Clearly, Case A applies if $q \in [0, \hat{q}]$ and Case B applies if $q \in (\hat{q}, 1]$. Notice that, intuitively, if the government is not very disposed toward the project ($T_{Gov} \simeq \frac{1}{2}$), then even a small bias in information is conducive to discipline ($\hat{q} \simeq 0$) since the government will quickly switch from w to n when signals are mixed.

Let $E(\pi^q_j)$ be the government's expected payoff if Case $j = A, B$ applies. Let $\Delta\pi^q_j \equiv E(\pi^q_j) - E(\pi^G_{nd})$. (The subscript nd is used to emphasize that the government picks w when the signals are mixed and is therefore not disciplined by public opinion.) In the Appendix we show that

$$\begin{aligned}\Delta\pi^q_A \;=\; & \underbrace{2\Pr\left((\alpha,\alpha)^q \mid (\varnothing,\alpha)^G\right)\Pr((\varnothing,\alpha)^G)}_{=qR}E \\ & - \underbrace{\Pr\left((\alpha,\alpha)^q \mid (\varnothing,\varnothing)^G\right)\Pr((\varnothing,\varnothing)^G)}_{=\frac{1}{2}q^2V}[C_w - B - \sigma_-] \\ & - \underbrace{2\Pr\left((\alpha,\varnothing)^q \mid (\varnothing,\varnothing)^G\right)\Pr((\varnothing,\varnothing)^G)}_{(1-q)qV}[T_{Gov} - \sigma_-]\,.\end{aligned} \tag{5}$$

Equation (5) can be easily interpreted. Suppose the government observes $(\alpha,\alpha)^q$ but the genuine signals are $(\varnothing,\alpha)^G$ or $(\alpha,\varnothing)^G$, which occurs with probability $2\Pr\left((\alpha,\alpha)^q \mid (\varnothing,\alpha)^G\right)\Pr((\varnothing,\alpha)^G) = qR$. In this case, biasing the screening technology is beneficial to the government for it can take the same decision as in the unbiased case without incurring electoral losses. The net gain is therefore E.[16] Similarly, suppose $(\alpha,\alpha)^q$ was generated by $(\varnothing,\varnothing)^G$. Now the government (erroneously) decides to implement the project. However, since the public supports the government, the net loss from switching from n to w (conditional on $(\varnothing,\varnothing)^G$) is not $[C_w - B + E - \sigma_-]$ but $[C_w - B - \sigma_-]$. Finally, it could be that mixed signals are observed, but the α-signal is fake. Biasing the screening technology hurts the government since the project is implemented when the status quo should have prevailed. The net loss is $T_{Gov} - \sigma_- = C_w - B + E - \sigma_-$.

Case B is similar, the only difference being that, when the signals are mixed, the likelihood that the α-signal is fake is so high that the government will choose n instead of w:

$$\begin{aligned}\Delta\pi_B^q \quad = \quad & \underbrace{2\Pr\left((\alpha,\alpha)^q \mid (\varnothing,\alpha)^G\right)\Pr((\varnothing,\alpha)^G)}_{=qR}E \\ & - \underbrace{\Pr\left((\alpha,\alpha)^q \mid (\varnothing,\varnothing)^G\right)\Pr((\varnothing,\varnothing)^G)}_{=\frac{1}{2}q^2V}[C_w - B - \sigma_-] \\ & - \underbrace{2\Pr\left((\alpha,\varnothing)^q \mid (\alpha,\varnothing)^G\right)\Pr((\alpha,\varnothing)^G)}_{(1-q)R}\left[\frac{1}{2} - T_{Gov}\right].\end{aligned} \tag{6}$$

Notice that when the signals are mixed and the α-signal turns out to be genuine, choosing the status quo is a mistake since $T_{Gov} \leq \frac{1}{2}$. Notice also that biasing the screening technology can be beneficial to the government ($\Delta\pi_A^q, \Delta\pi_B^q > 0$) only if electoral concerns are substantial.

4.2 Optimal Probability of Detection

We are now ready to tackle the problem of choosing the optimal size of the bias from the whole support $[0,1]$. Let $E(\pi_C^q)$ be the expected payoff to the government when $q \in (q^{\max}, 1]$, and $\Delta\pi_C^q \equiv E(\pi_C^q) - E(\pi_{nd}^G)$. For $T_{Gov} \leq \frac{1}{2}$, let $\Delta\pi_{TL}^q : [0,1] \to \mathbb{R}$ be defined as

$$\Delta\pi_{TL}^q = \begin{cases} \Delta\pi_A^q & \text{if } q \in [0, \min\{\hat{q}, q^{\max}\}] \\ \Delta\pi_B^q & \text{if } q \in (\hat{q}, q^{\max}] \text{ and } \hat{q} < q^{\max} \\ \Delta\pi_C^q & \text{if } q \in (q^{\max}, 1] \end{cases} . \tag{7}$$

[16]More formally, $\pi(\underbrace{d((\alpha,\alpha)^q)}_{w}, \underbrace{v((\alpha,\alpha)^q)}_{w} \mid (\varnothing,\alpha)^G) - \pi(\underbrace{d((\alpha,\varnothing)^G)}_{w}, \underbrace{v((\alpha,\varnothing)^G)}_{n} \mid (\varnothing,\alpha)^G) = E$, where $\pi(w, w \mid (\varnothing,\alpha)^G)$ is the actual payoff (conditional on the genuine signals) and $\pi(w, n \mid (\varnothing,\alpha)^G)$ is the counterfactual, had q been equal to zero.

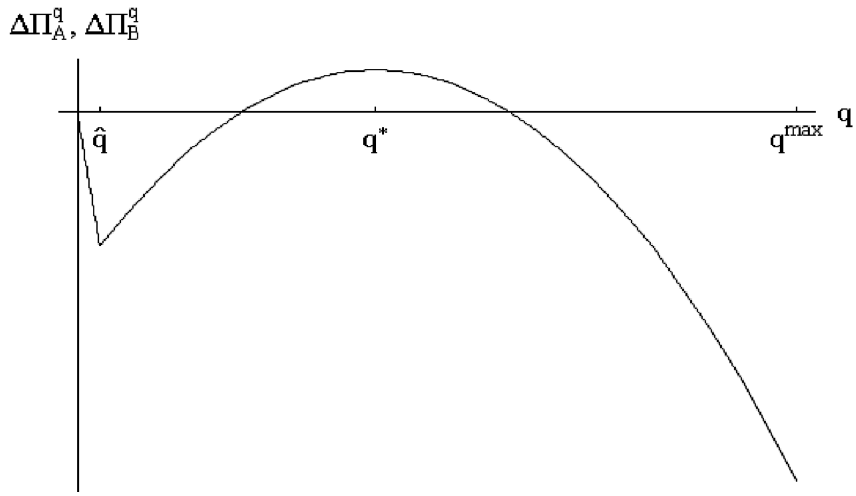

Figure 2: $\Delta\Pi_A^q$ and $\Delta\Pi_B^q$. Parameter values: $\theta = 4/5$, $C_w = 0.6$, $B = 0.27$, $E = 0.15$ ($\hat{q} \simeq 0.02$, $q^* \simeq 0.29$ and $q^{\max} \simeq 0.71$)

Notice that whenever $\Delta\pi_{TL}^q > 0$ for some q, the government will distort the screening technology. Furthermore, any q that maximizes the government's expected payoff will also maximize $\Delta\pi_{TL}^q$ since $E(\pi_{nd}^G)$ does not depend on q.

Proposition 3 *(i) $\Delta\pi_{TL}^q$ is strictly convex in q on $[0, \min\{\hat{q}, q^{\max}\}]$ and strictly concave in q on $(\hat{q}, q^{\max}]$.*[17] *Furthermore, $\Delta\pi_{TL}^q$ is continuous on $[0, q^{\max}]$ and $\left.\frac{\partial \Delta\pi_B^q}{\partial q}\right|_{q=\hat{q}<1} > 0$. (ii) $\Delta\pi_C^q < 0$ for all $q \in (q^{\max}, 1]$.*

Proposition 3(ii) shows that the government will not distort the screening technology so much that the public does not support w when $(\alpha, \alpha)^q$. Furthermore, since[18]

$$\arg\max_q \Delta\pi_B^q = r\frac{1/2 - (C_w - B)}{C_w - B - \sigma_-}$$

Proposition 3 implies that the optimum size of the bias will be either zero or $\min\left\{r\frac{1/2-(C_w-B)}{C_w-B-\sigma_-}, q^{\max}\right\}$ > 0, depending on parameter values.[19] Notice that if $q = 0$, the government will not be disciplined by public opinion. Thus, for large conflict of interest ($T_{Gov} \leq \frac{1}{2}$), the best outcome for the public–no bias in information and discipline–is not achievable.

[17] The second region may obviously not exist if $\hat{q} \geq q^{\max}$.

[18] The second order conditions are satisfied since $\frac{\partial^2 \Delta\pi_{TH}^q}{\partial q^2} = -V\left[C_w - B - \sigma_-\right] < 0$.

[19] In particular, $\arg\max \Delta\pi_{TL}^q = 0$ for E small enough. In the example of Figure 2, it suffices to take $E = 0.1$

4.3 Can the Public Benefit from Biased Information?

This subsection shows that, if the government is not very disposed toward the project and electoral concerns are small, the public will be able to achieve discipline with little distortion and will therefore prefer such situation to one in which no manipulations are possible.

Let $E(U_i^q)$, $i = A, B$, be the public's expected payoff when the size of the bias is q and $T_{Gov} \leq \frac{1}{2}$ (Case A and B). Let $E(U_{nd}^G)$ be the public's expected payoff when $q = 0$ and the government is *not* disciplined by public opinion (i.e., $d = w$ when signals are mixed). Define $\Delta U_i^q = E(U_i^q) - E(U_{nd}^G)$. It is not hard to show that[20]

$$\begin{aligned}
\Delta U_A^q &= -\underbrace{\Pr\left((\alpha,\alpha)^q \mid (\varnothing,\varnothing)^G\right)\Pr((\varnothing,\varnothing)^G)}_{=\frac{1}{2}q^2V}(C_w-\sigma_-) \\
&\qquad -\underbrace{2\Pr\left((\alpha,\varnothing)^q \mid (\varnothing,\varnothing)^G\right)\Pr((\varnothing,\varnothing)^G)}_{=(1-q)qV}(C_w-\sigma_-) \\
\Delta U_B^q &= -\underbrace{\Pr\left((\alpha,\alpha)^q \mid (\varnothing,\varnothing)^G\right)\Pr((\varnothing,\varnothing)^G)}_{=\frac{1}{2}q^2V}(C_w-\sigma_-) \\
&\qquad +\underbrace{2\Pr\left((\alpha,\varnothing)^q \mid (\alpha,\varnothing)^G\right)\Pr((\alpha,\varnothing)^G)}_{=(1-q)R}\left(C_w-\frac{1}{2}\right).
\end{aligned}$$

These equations can also be easily interpreted. Consider the first equation (case A). When the government distorts the screening technology, the public incurs two types of losses. First, the genuine signals might have been $(\varnothing,\varnothing)$, but mixed signals are observed. Secondly, the genuine signals might have been $(\varnothing,\varnothing)$, but (α,α) are observed. In both case, the government's decision changes from the n to w and the expected loss to the public is the difference between the payoffs in these two scenarios conditional on the genuine signals being $(\varnothing,\varnothing)$.[21] Notice that the public incurs no loss if the genuine signals are $(\alpha,\varnothing)$ but (α,α) is observed instead, since under both the efficient and the biased technology (in case A), the choice of the government is the same.

Case B is more interesting. As before the public incurs a loss of $-C_w + \sigma_-$ when the genuine

[20] The derivation of these and other equations is similar to that of $\Delta\pi_A^q$ and $\Delta\pi_B^q$ and is therefore omitted.

[21] More formally (and assuming that $(\alpha,\alpha)^q$ was observed) $u(\underbrace{d((\alpha,\alpha)^q)}_{w} \mid (\varnothing,\varnothing)^G) - u(\underbrace{d((\varnothing,\varnothing)^G)}_{n} \mid (\varnothing,\varnothing)^G) = -C_w + \sigma_-$ where $u(d \mid \mathbf{s}^G)$ is the expected payoff to the public if the government policy is d, conditional on the signals being $\mathbf{s}^G$.

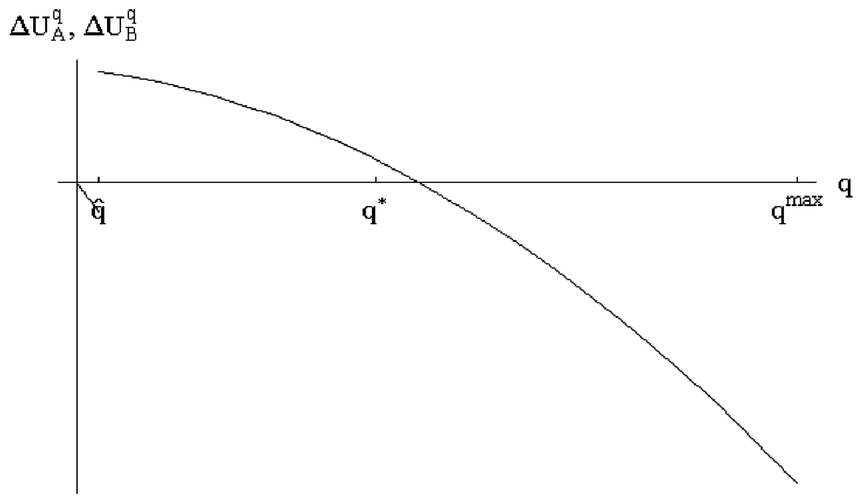

Figure 3: ΔU_A^q and ΔU_B^q for the same parameter values as in Figure 2

signals are $(\varnothing, \varnothing)$ but (α, α) is observed. Now however the government does not implement the project when the signals are mixed. This implies that the public does not incur any loss when instead of $(\varnothing, \varnothing)^G$ a mixed signal is observed since the government choice is the same. More importantly, the new situation can actually benefit the public. To see that, notice that if both the genuine and the observed signals are mixed (which occurs with probability $(1-q)R$), then a biased screening technology changes the decision of the government in the direction the public wants, namely from w to n.

The qualitative features of ΔU_B^q and ΔU_A^q are shown in Figure 3. Both functions are decreasing in q since, for any *given* decision rule, biased information hurts the public. However, when the screening technology is so biased that the government changes its decision rule, the public's payoff 'jumps up'. To see that, write

$$\Delta U_B^q - \Delta U_A^q = (1-q)\left[R\left(C_w - \frac{1}{2}\right) + qV\left(C_w - \sigma_-\right)\right] \geq 0. \tag{8}$$

This function measures the value of discipline for given q. Particularly interesting is the case where the size of the bias is small, since then $\Delta U_B^q - \Delta U_A^q \simeq R\left(C_w - \frac{1}{2}\right)$, which is just the probability that genuine signals are mixed (R) times the gain from switching from w to n, $C_w - \frac{1}{2}$. Importantly, since these gains are strictly positive, if the public can achieve discipline with little distortion, it will prefer such situation to one where no distortion at all is possible (i.e., $\Delta U_B^q > 0$

for q sufficiently small).

We now show that if the government is not very disposed toward the project, the equilibrium size of the bias can be made arbitrarily small. Suppose $T_{Gov} \simeq \frac{1}{2}$. Then even a very small bias in information induces the government to switch from w to n when the signals are mixed: $\hat{q} \simeq 0$. Furthermore $\arg\max_q \Delta\pi_B^q \simeq r \, \frac{E}{C_w - B - \sigma_-} \equiv q^*$. Notice that if E is small or signals are precise[22] q^*will be small and therefore $q^* < q^{\max}$ and $\Delta U_B^{q^*} > 0$. Furthermore, $\Delta\pi_B^{q^*} \simeq \frac{1}{2} q^* RE > 0$. Thus, as illustrated in Figure 2 and 3, when the government is not very disposed toward the project, it may well be the case that the public prefers to tolerate a small amount of bias (optimally chosen by the government) in order to more closely align the government's interests with its own.

4.4 $\mathbf{T}_{Gov} > \frac{1}{2}$

We now briefly consider the case where $T_{Gov} > \frac{1}{2}$. In this scenario, electoral concerns E are large relative to the private benefits B and the government is always disciplined by public opinion: $d = w$ if and only if $(\alpha, \alpha)^q$ (of course, in equilibrium $q \leq q^{\max}$). Let $E(\pi_d^q)$ be the expected payoff to the government when the size of the bias is q and the government is disciplined by public opinion. Let $\Delta\pi_{TH}^q \equiv E(\pi_d^q) - E(\pi_d^G)$, $q \in [0, q^{\max}]$:

$$\begin{aligned} \Delta\pi_{TH}^q \;=\;& \underbrace{2 \Pr\left((\alpha,\alpha)^q \mid (\varnothing,\alpha)^G\right) \Pr((\varnothing,\alpha)^G)}_{= qR} \left[\frac{1}{2} - (C_w - B)\right] \\ & - \underbrace{\Pr\left((\alpha,\alpha)^q \mid (\varnothing,\varnothing)^G\right) \Pr((\varnothing,\varnothing)^G)}_{= \frac{1}{2} q^2 V} \left[C_w - B - \sigma_-\right]. \end{aligned}$$

Let q^{TH} denote the optimum size of the bias. Notice that, for small private benefits ($C_w - B \geq \frac{1}{2}$), $\Delta\pi_{TH}^q < 0$. The following result obtains immediately.

Proposition 4 *Suppose $T_{Gov} > \frac{1}{2}$. The government is always disciplined by public opinion. Furthermore, if $C_w - B \geq \frac{1}{2}$, $q^{TH} = 0$. If $C_w - B < \frac{1}{2}$, $q^{TH} = \min\left\{ r \, \frac{1/2 - (C_w - B)}{C_w - B - \sigma_-}, \; q^{\max} \right\} > 0$.*

Clearly, $q^{TH} = 0$ coupled with discipline is the best outcome that the public can achieve. Notice also that q^{TH} hurts the public only when the genuine signals do not strongly support w but $(\alpha, \alpha)^q$ is observed. Indeed, those are the cases when the government policy switches from n

[22] Indeed, $r = \frac{2\theta(1-\theta)}{\theta^2 + (1-\theta)^2} \simeq 0$ if $\theta \simeq 1$.

to w. $\Delta U_{TH}^q = E(U_d^q) - E(U_d^G)$ is thus given by

$$\begin{aligned}\Delta U_{TH}^q &= -\underbrace{\Pr\left((\alpha,\alpha)^q \mid (\varnothing,\varnothing)^G\right)\Pr((\varnothing,\varnothing)^G)}_{=\frac{1}{2}q^2V}(C_w - \sigma_-)\\ &\qquad - \underbrace{2\Pr\left((\alpha,\alpha)^q \mid (\alpha,\varnothing)^G\right)\Pr((\alpha,\varnothing)^G)}_{=qR}\left(C_w - \frac{1}{2}\right).\end{aligned}$$

In the following, however, we will need to compare the present scenario ($E(U_d^q)$) to a situation where there is both no bias in information and no discipline: $E(U_{nd}^G)$. One might conjecture that $E(U_d^q) - E(U_{nd}^G) = \Delta U_{TH}^q + R(C_w - \frac{1}{2})$ since $R(C_w - \frac{1}{2})$ is the value of discipline at $q = 0$ (see equation (8)). That conjecture turns out to be correct:

$$\begin{aligned}E(U_d^q) - E(U_{nd}^G) &= -\underbrace{\Pr\left((\alpha,\alpha)^q \mid (\varnothing,\varnothing)^G\right)\Pr((\varnothing,\varnothing)^G)}_{=\frac{1}{2}q^2V}(C_w - \sigma_-)\\ &\qquad + \underbrace{2\Pr\left((\alpha,\varnothing)^q \mid (\alpha,\varnothing)^G\right)\Pr((\alpha,\varnothing)^G)}_{=(1-q)R}\left(C_w - \frac{1}{2}\right).\end{aligned}$$

5 Full Disclosure, q Unobservable

We now turn to the case where signals are disclosed but the public ignores the extent to which the process of information gathering and evaluation has been distorted. Under some conditions, the new game will have the same equilibria as the game where q is observable. However, when these conditions are not fulfilled, no pure strategy Nash equilibrium exists.

Suppose $T_{Gov} \leq \frac{1}{2}$. Define $\Delta\tilde{\pi}_{TL}^q : [0,1] \to \mathbb{R}$ as follows:

$$\Delta\tilde{\pi}_{TL}^q = \begin{cases} \Delta\pi_A^q & \text{if } q \in [0,\hat{q}] \\ \Delta\pi_B^q & \text{if } q \in (\hat{q},1] \end{cases}. \tag{9}$$

Notice that if citizens always support w when the signals are (α,α), this is precisely the function that the government should maximize. Indeed, equation (9) is the same as (7) provided $q^{\max} = 1$. Proposition 3 then makes clear that $\arg\max_{q\in[0,1]} \Delta\tilde{\pi}_{TL}^q$ will be either 0 or $\min\left\{r\frac{1/2-(C_w-B)}{C_w-B-\sigma_-}, 1\right\}$, depending on the parameters of the model. If there is more than one q that maximizes $\Delta\tilde{\pi}_{TL}^q$, we take the smallest.

Now suppose $T_{Gov} > \frac{1}{2}$. Let $\Delta\tilde{\pi}_{TH}^q = \Delta\pi_{TH}^q$ with $q \in [0,1]$ (not $[0, q^{\max}]$). Then

$$\arg\max_{q\in[0,1]} \Delta\tilde{\pi}_{TH}^q = \begin{cases} 0 & \text{if } C_w - B \geq \frac{1}{2} \\ \min\left\{r\frac{1/2-(C_w-B)}{C_w-B-\sigma_-}, 1\right\} & \text{if } C_w - B < \frac{1}{2} \end{cases}.$$

The main result of this section is the following.

Proposition 5 *Suppose $T_{Gov} \leq \frac{1}{2}$. A pure strategy Nash equilibrium (PSNE) of the full disclosure game with unobservable q exists if and only if $\arg\max_{q\in[0,1]} \Delta\tilde{\pi}_{TL}^q \leq q^{\max}$. If $T_{Gov} > \frac{1}{2}$, a PSNE exists if and only if $\arg\max_{q\in[0,1]} \Delta\tilde{\pi}_{TH}^q \leq q^{\max}$.*

When the above conditions are met, the analysis of the previous section is robust to the unobservability of q. It is not hard to see that such conditions are more likely to be fulfilled when the preferences of the public and those of the government are not too dissimilar. In fact, when the public is malleable (C_w close to $\frac{1}{2}$), $q^{\max}$ is close to one. The government has also less of an incentive to manipulate the information when B is small since $r\frac{1/2-(C_w-B)}{C_w-B-\sigma_-}$ is increasing in B. The corollary below follows immediately from Proposition 5 and the fact that $r\frac{1/2-(C_w-B)}{C_w-B-\sigma_-} \geq 1 \iff C_w - B \leq 1-\theta$.

Corollary 6 *Suppose $T_{Gov} > \frac{1}{2}$. (i) PSNE always exist if $C_w - B \geq \frac{1}{2}$. (ii) PSNE never exist if $C_w - B \leq 1-\theta$. (iii) If $C_w - B \in \left(1-\theta, \frac{1}{2}\right)$, PSNE may or may not exist depending on the value of C_w.*

The case when $T_{Gov} \leq \frac{1}{2}$ is less clear-cut. Nevertheless, PSNE always exist if E is small enough since then $\arg\max_{q\in[0,1]} \Delta\tilde{\pi}_{TL}^q = 0$. Notice that the example in Figure 2 is robust to the unobservability of q and $\arg\max_{q\in[0,1]} \Delta\tilde{\pi}_{TL}^q \simeq 0.29 > 0$.

6 Nondisclosure Scenario

In this section, the government commits not to disclose the signals and q is unobservable. This seems a reasonable set of assumptions: if no information is released, it cannot be easy for the media and the public to assess the reliability of the government's sources, for instance. However, the government policy is observable and this information can be used to update beliefs.

In this setting, mixed strategies are need. Let $(q, \gamma(\mathbf{s}^q))$ denote a strategy for the government, where $\gamma(\mathbf{s}^q)$ is a probability distribution over policies, conditional on $\mathbf{s}^q$. Let β denote a mixed strategy for the public. After observing policy $d \in \{w, n\}$, suppose the public believes the government is playing strategy $\left(\tilde{q}^d, \tilde{\gamma}^d\right)$. Let $\beta(d, \tilde{q}^d, \tilde{\gamma}^d)$ denote the public's optimal strategy given these beliefs.[23] Loosely speaking, a (perfect Bayesian) equilibrium of this game is a strategy profile $((q^*, \gamma^*), \beta^*)$ and a set of beliefs such that each player's strategy maximizes his payoff given his beliefs and the other player's strategy, and beliefs are formed according to $((q^*, \gamma^*), \beta^*)$ and Bayes' rule (in particular, $\left(\tilde{q}^d, \tilde{\gamma}^d\right) = (q^*, \gamma^*)$). If an action is not played in equilibrium, no restriction is imposed on beliefs.

Our first result shows that, so far as the quality of the information is concerned, secrecy is a good thing.

Proposition 7 *In the nondisclosure scenario, assume that after observing d the public holds beliefs $\left(\tilde{q}^d, \tilde{\gamma}^d\right)$. Let $\Pi(q)$ be the maximum expected payoff that the government can achieve in that case when the size of the bias is q. Then $\Pi(0) \geq \Pi(q)$.*

This proposition provides a justification for restricting attention to equilibria where $q = 0$. It says that for any given belief that the public may hold, knowing more ($q = 0$) cannot hurt the government. Indeed, as shown before, our notion of manipulation is equivalent to a garbling of information as defined by Marschak and Miyasawa (1968).

Two remarks are in order. The first observation is that, in interactive situations (games), it is well-known that more information is not always beneficial, not even for the player who holds it (see, e.g., Bassan, Scarsini, Zamir (1997)). However, as Kamien, Tauman and Zamir (1990) point out, "it is not the information itself that harms player 1 [the informed player] but the fact that player 2 knew that he had it" (p.133). In the nondisclosure scenario q is not publicly known. Of course, in equilibrium citizens will find out q, but for any *given* belief they might hold, the government has the incentive to acquire more information. It is this incentive that drives the result.

A second important remark is that Proposition 7 obviously relies on all screening technology

[23] This optimal strategy is obtained by first computing the probability that W is the true state of the world (say σ_P) given d, and then choosing w or n depending on whether $\sigma_P \gtrless C_w$ (a nondegenerate mixed strategy is played only if $\sigma_P = C_w$). Notice that in order to compute σ_P (via Bayes' rule), the public needs to have some beliefs about (q, γ).

being equally costly. If the unbiased screening technology ($q = 0$) was prohibitively costly, for instance, the result would not be true. Clearly, the present result is best interpreted as a comparison of gross expected payoffs.

We now characterize the equilibria of nondisclosure scenario for $q = 0$. Let $\hat{\sigma} = \Pr(W \mid w)$ be the probability that $S = W$ after observing $d = w$ and given that the government picks w with probability one when $\mathbf{s}^G \in \{(\alpha, \alpha), (\alpha, \varnothing), (\varnothing, \alpha)\}$ and n otherwise:

$$\hat{\sigma} = \frac{\theta^2 + R}{1 + R} \in \left(\frac{1}{2}, 1\right). \tag{10}$$

We write $\gamma(\mathbf{s}^q) = d$ when the government selects d with probability one and the signals are $\mathbf{s}^q$. Similarly, we write $\beta(d) = v$ when the public supports v with probability one and the government policy is d.

Proposition 8 *The following is a (perfect Bayesian) equilibrium of the nondisclosure game:* $q = 0$, $\gamma((\alpha,\alpha)^G) = w$, $\gamma((\varnothing,\varnothing)^G) = n$, $\beta(n) = n$. *Furthermore,* $\gamma((\alpha,\varnothing)^G) = \gamma((\varnothing,\alpha)^G)$ *and*

i) if $T_{Gov} \leq \frac{1}{2}$ *and* $\hat{\sigma} \geq (<) T_P$*, then* $\gamma((\alpha,\varnothing)^G) = w$ *and* $\beta(w) = w$ (n).

ii) if $T_{Gov} > \frac{1}{2}$, $C_w - B < \frac{1}{2}$ *and* $\hat{\sigma} \geq T_P$*, then* $\gamma((\alpha,\varnothing)^G) = w$ *and* $\beta(w) = w$.

iii) if $T_{Gov} > \frac{1}{2}$, $C_w - B < \frac{1}{2}$ *and* $\hat{\sigma} < T_P$*, then* $\gamma((\alpha,\varnothing)^G) = \begin{cases} w & \text{with probability } \check{p} \\ n & \text{with probability } 1-\check{p} \end{cases}$

and $\beta(w) = \begin{cases} w & \text{with probability } 1-\check{z} \\ n & \text{with probability } \check{z} \end{cases}$ *where* $\check{p} = \frac{\sigma_+^G - C_w}{2r(C_w - 1/2)}$ *and* $\check{z} = \frac{1/2 - (C_w - B)}{E}$.[24]

iv) if $C_w - B \geq \frac{1}{2}$*, then* $\gamma((\alpha,\varnothing)^G) = n$ *and* $\beta(w) = w$.

Not surprisingly, as the government becomes more disposed toward the project, the probability that the status quo is chosen when signals are mixed decreases. For large bias ($T_{Gov} \leq \frac{1}{2}$), this probability is zero and discipline is not attainable. For intermediate bias ($T_{Gov} > \frac{1}{2}$, $C_w - B \leq \frac{1}{2}$), the government selects the status quo with positive probability if the public is not malleable ($\hat{\sigma} < T_P$). We may label this outcome 'partial discipline'. However, when the public is malleable, the government 'gets away' with choosing $\gamma((\alpha,\varnothing)^G) = w$ because the public cannot distinguish whether w was induced by mixed signals or $(\alpha,\alpha)^G$ and since $\hat{\sigma} \geq T_P$, it ends up supporting the project. Finally, for small bias ($C_w - B > \frac{1}{2}$), both discipline and no manipulations obtain.

[24] It is not hard to see that $\hat{\sigma} < T_P \Longrightarrow \check{p} < 1$, and that $T_{Gov} > \frac{1}{2}, C_w - B < \frac{1}{2} \Longrightarrow \check{z} < 1$.

7 Transparency vs. Secrecy

This section briefly compares transparency (full disclosure) to secrecy (nondisclosure), from the public's viewpoint. In all our examples, we will restrict attention to situations which are robust to the unobservability of q.

Generally speaking, the choice between transparency and secrecy involves a basic tradeoff between the quality of information and discipline. Relative to full disclosure, a commitment not to disclose information is effective at protecting the integrity of the process of information gathering and evaluation. However, under full disclosure, the government is always disciplined by public opinion when the conflict of interest is not large ($T_{Gov} > \frac{1}{2}$) and sometimes even when it is large ($T_{Gov} \leq \frac{1}{2}$). When no information can be disclosed, instead, policymakers might not act according to public desires even for an 'intermediate' conflict of interest ($T_{Gov} > \frac{1}{2}$, $C_w - B \leq \frac{1}{2}$). Notice that such lack of discipline can be traced back to a lack of effective accountability since the public cannot distinguish whether the decision to implement the project was a good one ($\mathbf{s}^G = (\alpha, \alpha)$) or a bad one (e.g., $\mathbf{s}^G = (\alpha, \varnothing)$).

A few examples may help fix ideas. Suppose $T_{Gov} > \frac{1}{2}$ and $C_w - B < \frac{1}{2}$ (an 'intermediate' conflict of interest). Consider the case where $\hat{\sigma} \geq T_P$ (C_w small). The public's expected payoff in the full disclosure scenario is $E(U_d^q)$, $q = \min\left\{r \ \frac{1/2-(C_w-B)}{C_w-B-\sigma_-}, \ q^{\max}\right\} > 0$. $E(U_{nd}^G)$ is the public's expected payoff in the nondisclosure scenario (in fact, $q = 0$ and since $\hat{\sigma} \geq T_P$, the government is not disciplined by public opinion). Recall that

$$E(U_d^q) - E(U_{nd}^G) = -\frac{1}{2}q^2 V\left(C_w - \sigma_-\right) + (1-q)R\left(C_w - \frac{1}{2}\right).$$

Either transparency or secrecy can be optimal. Secrecy tends to perform well when discipline is not important ($C_w \simeq \frac{1}{2}$). In fact, if B is not 'too big', $q \simeq r\frac{B}{0.5-B-\sigma_-}$ ($q^{\max} \simeq 1$ since $C_w \simeq \frac{1}{2}$) and hence $E(U_d^q) < E(U_{nd}^G)$. By contrast, transparency tends to be preferable when the government is not very biased toward the project ($C_w - B \simeq \frac{1}{2}$) since then $q \simeq 0$ and $E(U_d^q) > E(U_{nd}^G)$.

Interestingly, the example in Figure 2 and 3 shows that transparency can be optimal even when disclosure per se (that is, without manipulation of information) would not be sufficient to discipline the government. In that example, discipline is sustained by the interplay between disclosure and manipulations and the public benefits from the fact that the government can manipulate the information. Notice that, should the government commit not to manipulate the information, the same outcome as in the nondisclosure scenario would prevail.

Finally, it is clear that when the conflict of interest is small ($C_w - B \geq \frac{1}{2}$), transparency and secrecy both achieve the best outcome (no distortion and discipline) and are therefore equivalent.

8 Extensions

8.1 Voluntary Disclosure

In the full disclosure scenario the government is assumed to share its information with the public. One interpretation is that the government *commits* to disclose its signals. However, in many situations decision makers have at least some discretion over whether or not to release information. Here we point out that allowing for voluntary disclosure would typically result in full disclosure since information is hard in this model, and the logic of the 'unraveling' result would apply (Milgrom (1981)). To see this, suppose q is observable and the public expects the government to always disclose favorable information (i.e., $(\alpha, \alpha)^q$, $q \leq q^{\max}$).[25] Lack of disclosure is interpreted in the most skeptical way, that is, as evidence that the signals did not strongly support implementation. The government has clearly a strict incentive to disclose favorable information. Whether or not the other signals are disclosed is inconsequential since the public will be able to infer that $\mathbf{s}^q \in \{(\alpha, \varnothing)^q, (\varnothing, \alpha)^q, (\varnothing, \varnothing)^q\}$ and will therefore support the status quo. Thus, this model is equivalent to the one studied in Section 4.

8.2 Direct Democracy

In the previous sections the decision maker had preferences that differed from those of the public. This need not always be the case. Increasingly often, in fact, decision-making powers are allocated to citizens via referendums or other forms of direct democracy (Matsusaka (2005)). This subsection briefly considers a variant of the model, termed direct democracy, where the decision is taken by the public itself after observing the signals. However, we retain the assumption that the government can manipulate the screening technology. For simplicity q is observable. q^{DD} denotes the optimal size of the bias.

[25] The analysis would probably extend straightforwardly to the case where q is not observable.

Proposition 9 *In the direct democracy scenario,* $q^{DD} \leq q^{\max}$ *and the project is implemented if and only if* $(\alpha, \alpha)^q$*. Furthermore,* $q^{DD} = 0$ *if* $C_w - B \geq \frac{1}{2}$ *and* $q^{DD} = \min\left\{r \ \frac{1/2-(C_w-B)}{C_w-B-\sigma_-},\ q^{\max}\right\}$ *if* $C_w - B < \frac{1}{2}$*.*

This result is similar to the one in Proposition 4, but now holds for $T_{Gov} \leq \frac{1}{2}$ as well. Notice that 'discipline' always obtains, but the screening technology can be distorted even when E is very small. A possible implication of this result is that an agent will always provide biased information to a principal that must ratify his decision, unless their preferences are congruent (i.e., $C_w - B \geq \frac{1}{2}$).

9 Concluding Remarks

We have studied the costs and benefits of disclosure of pre-decision information in a simple model where an agent (the government) cares about how its actions are perceived by a principal (the public) and can distort the information brought to bear on a decision. In line with conventional wisdom, transparency makes the government more accountable and hence more responsive to public desires. However, disclosure also creates incentives to distort the process of information gathering and evaluation. Surprisingly, we showed that manipulations may also have a bright side, for they can help more closely align the interests of the agent with those of the principal.

The present model is a very stylized one and we believe that many reasonable extensions would make transparency even more desirable. For instance, we assumed that the government can freely manipulate the process of information gathering and evaluation. This is obviously an extreme assumption. The extent to which bureaucracies can protect their independence from political interference is a crucial issue that we have completely neglected. Clearly, if reports could be completely insulated from partisan manipulation, a policy of full disclosure would dominate nondisclosure. Another important, implicit assumption we made was that different disclosure environments do not affect *per se* the quality of information. However, one of the main virtues of transparency is that independent analysts can verify government information and thinking about an issue. That would not only weed out forgeries, but would also minimize 'innocent' reporting and judgement errors. Finally, This paper posits that the government is not penalized just for biasing the information. Relaxing this assumption would reduce the incentives for manipulation and would make transparency more desirable.

Appendix: Proofs

Proof of Lemma 1. (i) Let $\sigma_{Gov} = \sigma_P = \sigma$. The public supports the decision to implement the project if $\Pr(W|\sigma) \geq C_w$. This implies that $\Pr(W|\sigma) \geq C_w - B + E(1_{\{w,w\}} - 1_{\{n,w\}}) = C_w - B - E$. Part (ii) follows immediately from (2). ■

Proof of Lemma 2. If $v = w$, the claim follows from Lemma 1. $v = n$ implies that $\Pr(W|I) < C_w$ and that the government will choose w if $\Pr(W|\sigma) \geq C_w - B + E$. This can never be the case if $E \geq B$. ■

Derivation of $\Delta\pi_i^q$, i=A,B

Recall that

$$E(\pi^q) = \sum_{\mathbf{s}^q \in \Omega^2} \Pr(\mathbf{s}^q)\pi(d(\mathbf{s}^q), v(\mathbf{s}^q) \mid \mathbf{s}^q).$$

We start with Case A and then consider Case B.

Case A:

$$\begin{aligned} E(\pi_A^q) &= \frac{1}{2}(\underbrace{[\theta + q(1-\theta)]^2 + [(1-\theta) + q\theta]^2}_{V+2qR+q^2V})\left[-C_w(1-\sigma_+^q) + B\right] \\ &\quad +2\frac{1}{2}(1-q)\underbrace{[(\theta + q(1-\theta))(1-\theta) + ((1-\theta) + q\theta)\theta]}_{R+qV}[-C_w(1-\sigma^q) + B + E] \\ &\quad +\frac{1}{2}(1-q)^2 V\,[-C_n\sigma_-] \\ &= \frac{1}{2}[V + 2qR + q^2V]\,[-C_w + B] + \frac{1}{2}\left[\theta^2 + q^2(1-\theta)^2 + qR\right] C_w \\ &\quad +(1-q)\,[R + qV]\,[-C_w + B + E] + (1-q)\left[R/2 + q(1-\theta)^2\right] C_w \\ &\quad +\frac{1}{2}(1-q)^2 V\,[-C_n\sigma_-] \end{aligned} \tag{11}$$

where $\Pr((\alpha,\alpha)^q) = \frac{1}{2}([\theta + q(1-\theta)]^2 + [(1-\theta) + q\theta]^2)$, $\Pr((\alpha,\varnothing)^q) = \Pr((\varnothing,\alpha)^q) = \frac{1}{2}(1-q)\,[(\theta + q(1-\theta))(1-\theta) + ((1-\theta) + q\theta)\theta]$, $\Pr((\varnothing,\varnothing)^q) = \frac{1}{2}(1-q)^2 V$. The first line of (11) can be written as

$$\frac{1}{2}V\underbrace{\left[-C_w + B + \frac{\theta^2}{V}C_w\right]}_{\pi(w,w|(\alpha,\alpha)^G)} + qR\underbrace{\left[-C_w + B + \frac{1}{2}C_w\right]}_{\pi(w,w|(\alpha,\varnothing)^G)} + \frac{1}{2}q^2V\underbrace{\left[-C_w + B + \frac{(1-\theta)^2}{V}C_w\right]}_{\pi(w,w|(\varnothing,\varnothing)^G)}.$$

Similarly, the second line of (11) can be written as

$$(1-q)R\underbrace{\left[-C_w+B-E+\frac{1}{2}C_w\right]}_{\pi(w,n|(\alpha,\varnothing)^G)}+q(1-q)V\underbrace{\left[-C_w+B-E+\frac{(1-\theta)^2}{V}C_w\right]}_{\pi(w,n|(\varnothing,\varnothing)^G)}$$

and the third as

$$\frac{1}{2}(1-q)^2V\underbrace{[-C_n\sigma_-]}_{\pi(n,n|(\varnothing,\varnothing)^G)}.$$

Since $E(\pi^G)=\frac{1}{2}V\pi(w,w|(\alpha,\alpha)^G)+R\pi(w,n|(\alpha,\varnothing)^G)+\frac{1}{2}V\pi(n,n|(\varnothing,\varnothing)^G)$,

$$\begin{aligned}\Delta\pi_A^q &= qR\underbrace{\left[\pi(w,w|(\alpha,\varnothing)^G)-\pi(w,n|(\alpha,\varnothing)^G)\right]}_{E}\\ &\quad+\frac{1}{2}q^2V\underbrace{\left[\pi(w,w|(\varnothing,\varnothing)^G)-\pi(n,n|(\varnothing,\varnothing)^G)\right]}_{-C_w+B+\sigma_-}\\ &\quad+q(1-q)V\underbrace{\left[\pi(w,n|(\varnothing,\varnothing)^G)-\pi(n,n|(\varnothing,\varnothing)^G)\right]}_{-C_w+B-E+\sigma_-}.\end{aligned}$$

Case B. This case is very similar to the previous one, the only difference being that when the signals are mixed, the government now chooses n, not w. The second line in (11) now reads

$$\begin{aligned}2\frac{1}{2}(1-q)(R+qV)\underbrace{[-C_n\sigma^q]}_{\pi(n,n|(\alpha,\varnothing)^q)} &= (1-q)\left[\frac{R}{2}+q(1-\theta)^2\right][-C_n]\\ &= (1-q)R\underbrace{\left[-\frac{1}{2}C_n\right]}_{\pi(n,n|(\alpha,\varnothing)^G)}+q(1-q)V\underbrace{[-C_n\sigma_-]}_{\pi(n,n|(\varnothing,\varnothing)^G)}\end{aligned}$$

since $\sigma^q=\frac{\frac{R}{2}+q(1-\theta)^2}{R+qV}$. Thus

$$\begin{aligned}E(\pi_B^q) &= \frac{1}{2}V\pi(w,w|(\alpha,\alpha)^G)+qR\pi(w,w|(\alpha,\varnothing)^G)+\frac{1}{2}q^2V\pi(w,w|(\varnothing,\varnothing)^G)\\ &\quad+(1-q)R\pi(n,n|(\alpha,\varnothing)^G)+q(1-q)V\pi(n,n|(\varnothing,\varnothing)^G)\\ &\quad+\frac{1}{2}(1-q)^2V\pi(n,n|(\varnothing,\varnothing)^G)\end{aligned}$$

and therefore $\Delta\pi_B^q$ is as given in (6).

Proof of Proposition 3. Differentiating $\Delta\pi_A^q$ twice with respect to q yields

$$\begin{aligned}\frac{\partial \Delta\pi_A^q}{\partial q} &= RE - qV(C_w - B - \sigma_-) - (1-2q)V\left[T_{Gov} - \sigma_-\right] \\ \frac{\partial^2 \Delta\pi_A^q}{\partial q^2} &= V(T_{Gov} - \sigma_- + E) > 0.\end{aligned}$$

Similarly, differentiating $\Delta\pi_B$ twice with respect to q yields

$$\begin{aligned}\frac{\partial \Delta\pi_B^q}{\partial q} &= RE - qV(C_w - B - \sigma_-) + R\left[\frac{1}{2} - T_{Gov}\right] \\ \frac{\partial^2 \Delta\pi_B^q}{\partial q^2} &= -V(C_w - B - \sigma_-) < 0.\end{aligned}$$

Thus $\Delta\pi_A^q$ ($\Delta\pi_B^q$) is strictly convex (concave) in q. Notice that

$$\left.\frac{\partial \Delta\pi_B^q}{\partial q}\right|_{q=\hat{q}} = RE\left(1 + \frac{\frac{1}{2} - T_{Gov}}{T_{Gov} - \sigma_-}\right) > 0, \quad \hat{q} < 1.$$

Continuity of $\boldsymbol{\Delta\pi_{TL}^q}$ on $[0, q^{\max}]$ (and of $\Delta\tilde{\pi}_{TL}^q$ on $[0,1]$). Suppose that $\hat{q} < q^{\max}$. $\Delta\pi_A^q$ and $\Delta\pi_B^q$ are continuous in q on their respective domains. Thus, we just need to show that $\Delta\pi_A^q = \Delta\pi_B^q$ at $\hat{q}$. Recall that $\hat{q}$ solves $\sigma^q = \frac{\frac{R}{2} + q(1-\theta)^2}{R + qV} = T_{Gov}$ if $T_{Gov} \geq 1 - \theta$. Solving that equation yields

$$\hat{q} = \frac{\left(\frac{1}{2} - T_{Gov}\right) R}{VT_{Gov} - (1-\theta)^2} = r\ \frac{1/2 - T_{Gov}}{T_{Gov} - \sigma_-}. \tag{12}$$

Notice that $\Delta\pi_A^{\hat{q}} = \Delta\pi_B^{\hat{q}}$ is equivalent to $-(1-\hat{q})\,\hat{q}V\left[T_{Gov} - \sigma_-\right] = -(1-\hat{q})\,R\left[\frac{1}{2} - T_{Gov}\right]$

$$\Leftrightarrow \quad \frac{-\left(\frac{1}{2} - T_{Gov}\right) R}{VT_{Gov} - (1-\theta)^2} V\left[T_{Gov} - \frac{(1-\theta)^2}{V}\right] = R\left(T_{Gov} - \frac{1}{2}\right)$$

which is clearly true.

It remains to be shown that $\Delta\pi_C^q < 0$ for all $q \in (q^{\max}, 1]$. Suppose q is sufficiently low that, if signals are mixed, the government chooses w, and denote this case C/A. Then

$$\begin{aligned}&\Delta\pi_{C/A}^q = \\ &\quad \tfrac{1}{2}V\left[-C_w + B - E + C_w\sigma_+^G\right] + qR\left[-C_w + B - E + \tfrac{1}{2}C_w\right] + \tfrac{1}{2}q^2V\left[-C_w + B - E + C_w\sigma_-\right] \\ &\quad +(1-q)R\left[-C_w + B - E + \tfrac{1}{2}C_w\right] + q(1-q)V\left[-C_w + B - E + C_w\sigma_-\right] + \tfrac{1}{2}(1-q)^2V\left[-C_n\sigma_-\right]\end{aligned}$$

$-\left\{\frac{1}{2}V\left[-C_w+B+C_w\sigma_+^G\right]+R\left[-C_w+B-E+\frac{1}{2}C_w\right]+\frac{1}{2}V\left[-C_n\sigma_-\right]\right\}$

$=-\frac{1}{2}VE-q\left(1-\frac{q}{2}\right)V\left[C_w-B+E-\sigma_-\right]<0.$

The case where n is chosen when the signals are mixed is similar. ■

Proof of Proposition 5. Let $\tilde{q}^{TL}=\arg\max_{q\in[0,1]}\Delta\tilde{\pi}^q_{TL}$ (if there are multiple solutions, take the smallest). Suppose the public believes $q=q^e\leq q^{\max}$. Since $q^e\leq q^{\max}$, the public supports w when (α,α) are observed. The government will therefore $\max_{q\in[0,1]}\Delta\tilde{\pi}^q_{TL}$ and choose $\tilde{q}^{TL}$. Consistency requires $q^e=\tilde{q}^{TL}$. Since we assumed $q^e\leq q^{\max}$, we have a PSNE provided $\tilde{q}^{TL}\leq q^{\max}$.

This reasoning fails if $\tilde{q}^{TL}>q^{\max}$. In that case, no PSNE exists. Indeed, if $q^e\leq q^{\max}$, beliefs are not consistent with the play of the game: $q^e\neq\tilde{q}^{TL}$. Notice that the belief $q^e>q^{\max}$ cannot be supported in equilibrium. In fact, if $q^e>q^{\max}$, the public never supports w. The government's best response is then to set $q=0$ (this follows from Proposition 3(ii)). Again, beliefs are inconsistent with the play of the game since $q=0$ but $q^e>q^{\max}$.

The proof for $T_{Gov}>\frac{1}{2}$ is similar. ■

Proof of Proposition 7. After observing d, suppose the public believes that the government plays strategy $\left(\tilde{q}^d,\tilde{\gamma}^d\right)$. Let $\beta(d,\tilde{q}^d,\tilde{\gamma}^d)$ denote the public's optimal mixed strategy, given these beliefs. Suppose the government actually picks $q>0$. (Since $\tilde{q}^d$ may or may not be equal to q, this need not be an equilibrium.) The best the government can achieve is

$$\begin{aligned}\Pi(q) &= \sum_{\mathbf{s}^q\in\Omega^2}\Pr(\mathbf{s}^q)\max_d\sum_{S\in\Lambda}\pi\left(d,\beta(d,\tilde{q}^d,\tilde{\gamma}^d)\mid S\right)\Pr(S\mid\mathbf{s}^q)\\ &\equiv \sum_{\mathbf{s}^q\in\Omega^2}\Pr(\mathbf{s}^q)\sum_{S\in\Lambda}\pi\left(\gamma^*(\mathbf{s}^q),\beta(d,\tilde{q}^d,\tilde{\gamma}^d)\mid S\right)\Pr(S\mid\mathbf{s}^q).\end{aligned}$$

Notice that

$$\begin{aligned}\Pr(\mathbf{s}^q)\Pr(S\mid\mathbf{s}^q) &= \Pr(\mathbf{s}^q\mid S)\Pr(S)\\ &= \sum_{\mathbf{s}^G\in\Omega^2}\Pr(\mathbf{s}^q\mid\mathbf{s}^G)\Pr\left(\mathbf{s}^G\mid S\right)\Pr(S)\\ &= \sum_{\mathbf{s}^G\in\Omega^2}\Pr(\mathbf{s}^q\mid\mathbf{s}^G)\Pr\left(S\mid\mathbf{s}^G\right)\Pr(\mathbf{s}^G)\end{aligned}$$

where the first and third equality follow from Bayes' rule and the second equality follows from

$\Pr(\mathbf{s}^q \mid S) = \sum_{\mathbf{s}^G \in \Omega^2} \Pr(\mathbf{s}^q, \mathbf{s}^G \mid S)$ and the 'garbling' condition (3). Thus

$$\Pi(q) = \sum_{\mathbf{s}^G \in \Omega^2} \Pr(\mathbf{s}^G) \sum_{\mathbf{s}^q \in \Omega^2} \Pr(\mathbf{s}^q \mid \mathbf{s}^G) \sum_{S \in \Lambda} \pi \left(\gamma^*(\mathbf{s}^q), \beta(d, \tilde{q}^d, \tilde{\gamma}^d) \mid S \right) \Pr\left(S \mid \mathbf{s}^G\right).$$

Now suppose the government picks $q = 0$. (Since only d is observable, the public's beliefs are still given by $(\tilde{q}^d, \tilde{\gamma}^d)$ and the public's best reply is β.) The government's maximum expected payoff is

$$\begin{aligned} \Pi(0) &= \sum_{\mathbf{s}^G \in \Omega^2} \Pr(\mathbf{s}^G) \max_d \sum_{S \in \Lambda} \pi \left(d, \beta(d, \tilde{q}^d, \tilde{\gamma}^d) \mid S \right) \Pr(S \mid \mathbf{s}^G) \\ &\equiv \sum_{\mathbf{s}^G \in \Omega^2} \Pr(\mathbf{s}^G) \sum_{S \in \Lambda} \pi \left(\gamma^*(\mathbf{s}^G), \beta(d, \tilde{q}^d, \tilde{\gamma}^d) \mid S \right) \Pr(S \mid \mathbf{s}^G). \end{aligned}$$

From

$$\textstyle\sum_{S \in \Lambda} \pi \left(\gamma^*(\mathbf{s}^G), \beta(d, \tilde{q}^d, \tilde{\gamma}^d) \mid S \right) \Pr(S \mid \mathbf{s}^G) \geq \sum_{S \in \Lambda} \pi \left(\gamma^*(\mathbf{s}^q), \beta(d, \tilde{q}^d, \tilde{\gamma}^d) \mid S \right) \Pr\left(S \mid \mathbf{s}^G\right)$$

and

$$\textstyle\sum_{\mathbf{s}^q \in \Omega^2} \Pr(\mathbf{s}^q \mid \mathbf{s}^G) = 1 \qquad \forall \mathbf{s}^G$$

it easily follows that $\Pi(0) \geq \Pi(q)$. ■

Proof of Proposition 8. By Proposition 7, $q = 0$ maximizes the government's expected payoff for any strategy that the public might play. It is straightforward to check that the remaining strategies form an equilibrium. For the sake of brevity, we will focus on part (iii) (mixed strategies). Suppose $T_{Gov} > \frac{1}{2}$, $C_w - B < \frac{1}{2}$ and $\hat{\sigma} < T_P$. Clearly $\gamma((\varnothing, \varnothing)^G) = n$ since

$$-C_n \sigma_- > -C_w(1 - \sigma_-) + B - \check{z}E \iff C_w - B > \sigma_- - \check{z}E.$$

It is also not hard to see that $\beta(n) = n$. Now notice that if $\gamma((\alpha, \varnothing)^G) = w$ with probability one then $\beta(w) = n$ since $\hat{\sigma} < T_P$. But then $\gamma((\alpha, \varnothing)^G) = w$ cannot be optimal since $T_{Gov} > \frac{1}{2}$. Similarly, if $\gamma((\alpha, \varnothing)^G) = n$ with probability one then $\beta(w) = w$ since $\sigma_+^G > T_P$. But then $\gamma((\alpha, \varnothing)^G) = n$ is not optimal since $C_w - B < \frac{1}{2}$. We therefore need mixed strategies. The indifference condition for the government is

$$-C_n \frac{1}{2} = -C_w \left(1 - \frac{1}{2}\right) + B - \check{z}E \implies \check{z} = \frac{1/2 - (C_w - B)}{E}.$$

The indifference condition for the public is $\Pr(W \mid w, \check{p}) = C_w$, where $\Pr(W \mid w, \check{p})$ is the probability that $S = W$ after observing $d = w$ given that the government plays w with probability one if $(\alpha, \alpha)^G$, with probability $\check{p}$ if the signals are mixed and with probability zero if $(\varnothing, \varnothing)^G$. Using Bayes' rule

$$\Pr(W \mid w, \check{p}) = \frac{\theta^2 + \check{p}R}{V + 2\check{p}R} = C_w \implies \check{p} = \frac{\theta^2 - C_w V}{R\,(2C_w - 1)} = \frac{\sigma_+^G - C_w}{2r\,(C_w - 1/2)}.$$

To conclude the proof, notice that if the government plays w with positive probability when signals are mixed, it must play w with probability one when $(\alpha, \alpha)^G$. ■

Proof of Proposition 9. Clearly in equilibrium $q \in [0, q^{\max}]$ otherwise the public would never implement the project. Notice that the public will choose the status quo for all q and any signal except $(\alpha, \alpha)^q$ since $\Pr(W|(\varnothing, \varnothing)^q) \leq \Pr(W|(\alpha, \varnothing)^q) \leq \frac{1}{2} < T_P$. Let $E(\pi_{DD}^q)$ be the government's expected payoff when the size of the bias is q in the direct democracy scenario. Clearly if the policy is chosen by the public, the government never pays electoral costs. Let $\Delta\pi_{DD}^q = E(\pi_{DD}^q) - E(\pi_{DD}^0)$. Then

$$\begin{aligned}
\Delta\pi_{DD}^q \;=\; & \frac{1}{2}V\left[-C_w + B + C_w\sigma_+^G\right] + qR\left[-C_w + B + \frac{1}{2}C_w\right] + \frac{1}{2}q^2V\left[-C_w + B + C_w\sigma_-\right] \\
& +(1-q)R\left[-\frac{1}{2}C_n\right] + q(1-q)V\left[-C_n\sigma_-\right] + \frac{1}{2}(1-q)^2V\left[-C_n\sigma_-\right] \\
& - \left\{\frac{1}{2}V\left[-C_w + B + C_w\sigma_+^G\right] + R\left[-\frac{1}{2}C_n\right] + \frac{1}{2}V\left[-C_n\sigma_-\right]\right\}.
\end{aligned}$$

This expression yields

$$\Delta\pi_{DD}^q = qR\left[-C_w + B + \frac{1}{2}\right] + \frac{1}{2}q^2V\underbrace{(-C_w + B + \sigma_-)}_{<0}.$$

Clearly, $q^{DD} = 0$ if $C_w - B \geq \frac{1}{2}$ and $q^{DD} = \min\left\{r\ \frac{1/2-(C_w-B)}{C_w-B-\sigma_-},\ q^{\max}\right\} > 0$ if $C_w - B < \frac{1}{2}$. ■

References

[1] Arnold, D. R. (1987). 'Political Control of Administrative Officials', *Journal of Law, Economics and Organization*, vol. 3(2) (Fall), pp. 279-286.

[2] Bassan, B., Scarsini, M. and Zamir, S. (1997). '"I Don't Want to Know": Can It Be Rational?', mimeo.

[3] BMA survey of A&E waiting times (2005), downloadable at http://www.bma.org.uk /ap.nsf/Content/AandEwaiting.

[4] The Commission on the Intelligence Capabilities of the United States Regarding Weapons of Mass Destruction (2005) 'Report of the Commission on the Intelligence Capabilities of the United States Regarding Weapons of Mass Destruction', downloadable at http://www.wmd.gov/report/wmd_report.pdf

[5] Committee of Privy Counsellors. Chairman: The Rt Hon The Lord Butler of Brockwell KG GCB CVO (2004). 'Review of Intelligence on Weapons of Mass Destruction', downloadable at http://www.butlerreview.org.uk/.

[6] Cremer, J. (1995). 'Arm's Length Relationships', *Quarterly Journal of Economics*, vol.110(2) (May), pp.275-295.

[7] Dewatripont, M. and Tirole J. (1999). 'Advocates', *Journal of Political Economy*, vol.107(1), pp.1-39.

[8] The Economist (2004). 'Special Report Intelligence Failures', (July 17th) pp.23-25.

[9] The Economist (2005). 'Government Figures', (March 26th), pp.31-32.

[10] Frankel, M. (2001). 'Freedom of Information: Some International Characteristics', Working Paper, The Campaign for Freedom of Information.

[11] Garicano, L. and Posner R. A. (2005). 'Intelligence Failures: An Organizational Economics Perspective', forthcoming *Journal of Economic Perspectives*.

[12] Holmstrom, B. (1979). 'Moral Hazard and Observability', *Bell Journal of Economics*, vol.10 (Spring), pp.74-91.

[13] Holmstrom, B. (1999). 'Managerial Incentive Problems: A Dynamic Perspective', *Review of Economic Studies*, vol.66(1) (January), pp.169-182.

[14] Kamien, M. I., Tauman, Y. and Zamir S. (1990). 'On the Value of Information in Strategic Conflict', *Games and Economic Behavior*, vol.2, pp. 129-153.

[15] Krueger, A. B. and Laitin D. D. (2004). '"Misunderestimating" Terrorism', *Foreign Affairs*, vol.83(5) (September/October), pp.8-13.

[16] Jones, F. and Jones A.-M. (2005). 'Public Confidence in Official Statistics: An analysis based on data collected in the National Statistics Omnibus Survey', downloadable at http://www.statistics.gov.uk/about/data/public_confidence/reports.asp

[17] Lerner, J. and Tirole J. (2006). 'A Model of Forum Shopping, with Special Reference to Standard Setting Organizations', *American Economic Review*, forthcoming.

[18] Levy, G. (2005). 'Decision Making in Committees: Transparency, Reputation and Voting Rules', Working Paper.

[19] Maggi G. and Rodriguez-Clare A. (1995). 'Costly Distortion of Information in Agency Problems', *Rand Journal of Economics*, vol.26(4) (Winter), pp.675-689.

[20] Marschak, J. and Miyasawa K. (1968). 'Economic Comparability of Information Systems', *International Economic Review*, vol.9(2) (June), pp.137-174.

[21] Maskin, E. and Tirole J. (2004). 'The Politician and the Judge', *American Economic Review*, vol. 94(4) (September), pp.1034-1051.

[22] Matsusaka, J. G. (2005). 'Direct Democracy Works', *Journal of Economic Perspectives*, vol.19(2) (Spring), pp.185-206.

[23] Milgrom, P. (1981). 'Good News and Bad News: Representation Theorems and Applications', *Bell Journal of Economics*, vol.12(2) (Autumn), pp.380-391.

[24] Mullainathan, S. and Shleifer A. (2005). 'The Market for News', forthcoming *American Economic Review*.

[25] Persson T. and Tabellini G. (2002). *Political Economics. Explaining Economic Policy*, The MIT Press, Cambridge Massachusetts.

[26] Prat, A. (2005). 'The Wrong Kind of Transparency', *American Economic Review*, vol. 104(3) (June), pp. 946-59.

[27] Prat, A. (2006). 'The More Closely we are Watched, the Better we Behave?' in *Transparency, The Word and the Doctrines*, C. Hood and D. Heald eds, British Academy, forthcoming.

[28] Prendergast, C. (1993). 'A Theory of "Yes Men"', *American Economic Review*, vol. 83(4) (September), pp. 757-770.

[29] Ridley, F. F. (1991). 'Using Power To Keep Power – The need for Constitutional Checks', *Parliamentary Affairs*, vol. 44(4) (July), pp. 442-452.

[30] Roberts, A. (2006). 'Dashed Expectations: Governmental Adaptation to Transparency Rules' in *Transparency, The Word and the Doctrines*, C. Hood and D. Heald eds, British Academy, forthcoming.

[31] Select Committee on Intelligence United States Senate (2004). 'Report of the Select Committee on Intelligence on the U.S. Intelligence Community's Prewar Intelligence Assessments on Iraq', downloadable at http://www.gpoaccess.gov/serialset/creports/iraq.html

[32] Select Committee on Intelligence United States Senate (2004). 'Report on the U.S. Intelligence Community's Prewar Intelligence Assessments on Iraq. Conclusions' (Excerpted From Full Report). Downloadable at http://intelligence.senate.gov/

[33] Verrecchia, R. E. (2001). 'Essays on Disclosure', *Journal of Accounting and Economics*, vol.32(1-3) (December), pp.97-180.