University College London

Doctoral Thesis

Essays on Information and Political Economy

Author:
Boriss Ginzburgs

Supervisor:
Professor V Bhaskar

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

in the

Department of Economics

August 2015

Declaration of Authorship

I, Boris Ginzburg, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the work.

Signed:

Date:

Abstract

This thesis examines the role of information acquisition and information revelation in political and social interactions. Chapter 2 looks at a committee that needs to make a collective decision which gives every member a private state-dependent payoff. The committee can vote to learn the state at no cost. The chapter finds that the committee chooses to be uninformed whenever preferences of its members are sufficiently heterogeneous. Furthermore, members whose preferred state is revealed slower are better off. Chapter 3 analyses voting by a heterogeneous group in presence of private information. Unlike much of the literature, which looks at groups with relatively similar preferences and finds that voters generally do not vote according to their signals, Chapter 3 shows that sincere voting is an equilibrium strategy of all voters when preferences are sufficiently diverse. Chapter 4 looks at a government who can decide which news to censor and which news to reveal to a heterogeneous population of citizens. It turns out that stricter censorship is optimal when citizens are, overall, more supportive of the government. Finally, Chapter 5 looks at costly information disclosure by contestants competing for a prize. The chapter shows that an increase in competition leads to more information disclosure if and only if the cost of disclosure is high. Furthermore, when information is revealed with some noise, contestants are more likely to reveal it.

Acknowledgements

I would like to thank my supervisor, Professor V Bhaskar, whose invaluable support made this thesis possible. I am also grateful to Antonio Cabrales, Amanda Friedenberg, Philippe Jehiel, Emir Kamenica, Suehyun Kwon, Stephen Morris, Santiago Oliveros, and Nikita Rokeskiy.

Many of my fellow PhD students have given me invaluable help at various stages of this work. I want to give special thanks to Lena Aristodemou, Valerio Dotti, Joseph-Simon Gorlach, Michael Graber, Jose Alberto Guerra, Laura Jaitman, Christian Krestel, Warn Nuarpear Lekfuangfu, Myra Mohnen, Nicolas Motz, Sandra Polania Reyes, Cristina Rosemberg, Alireza Sepahsalari, Andreas Uthemann, Tanya Wilson.

Contents

List of Figures

Chapter 1

Introduction

This thesis explores the role of information in societies. In particular, it looks at the effect of information on political processes.

The second chapter of this thesis examines the decision of a group, such as a parliament or an electorate, to acquire information about a decision on which it needs to vote. Many decisions, such as economic reforms, education policies, or allocation of research funding, have consequences that are uncertain at the time when the decision is being made. Often, however, the group can vote to acquire information about the consequences before making the decision - for example, by voting to delay the decision until more information arrives. Yet often these opportunities are not pursued, and the group chooses to remain uninformed.

Chapter 2 analyses this phenomenon by modeling a committee that needs to vote on a reform which will give every member a private state-dependent payoff. The committee can vote to delay the decision. While the decision is being delayed, information can arrive and reveal the state. It turns out that even though delaying the decision is costless, not doing so (and thus voting "in ignorance") may be a rational collective decision. This happens when individual preferences are sufficiently heterogeneous. Hence, decisions on divisive issues are likely to be made with less consideration, and diverse societies are likely to adopt reforms in haste and to have less public debate. The chapter also finds that voters whose preferred state is revealed slower are better off. Finally, the chapter examines the question of what voting rule is optimal in situations when the committee has a choice of how much to learn. Under certain restrictions, it turns out that the optimal voting rule is a supermajority rule, and larger supermajority is optimal when gains and losses from the reform are more asymmetric or more random.

Chapter 3 continues to analyse the role of information in collective decisions, but does so from a different angle. Instead of explicitly looking at a collective decision to acquire information, it deals with situations in which information is dispersed among voters. Specifically, it models a setting in which a group of voters need to decide whether to approve a reform. A voter can either gain or lose from the reform, depending on a state of the world. Each voter receives a private signal about the state.

Prior research on models of this kind has mostly looked at settings in which voters have broadly similar interests - i.e. that the payoff difference from voting for the reform is monotone in the state across voters. The usual result is that in these situations, voters, at an equilibrium, do not in general vote according to their signals. This is a surprising result that does not have much empirical support.

Unlike much of the previous research, Chapter 3 considers a setup in which some voters gain from the reform in the first state, while others - in the second. It follows that when preferences are sufficiently heterogeneous - that is, when the minority is sufficiently large - there exists an equilibrium in which all voters act according to their signals.

Chapter 4 also looks at the role of information in collective decisions. But instead of examining a group's decision whether to seek information, it looks at a decision of an agent on whether to provide information to a group. The specific application is political censorship. The chapter looks at government that can choose what news to censor and what news to reveal to a population of voters. Voters differ in their attitudes towards the government. Upon hearing or not hearing the news, voters decide whether to support the government. The government is interested in maximising its share of support. The chapter finds that a shift in the number of extremists that is favourable from the government's point of view leads it to censor a smaller set of news. On the other hand, a favourable change in the attitudes of the population as a whole causes the government to censor a larger set of news. Thus, popular authoritarian regimes are likely to have stronger censorship, and a fall in popularity can induce a regime to relax media controls.

Chapter 5 similarly examines information disclosure. But unlike Chapter 4, it looks at information disclosure by competing agents. Specifically, Chapter 5 models a group of contestants are competing for a prize. Each contestant has a type, which is his private information. The prize is allocated by a decision-maker, who would like to give it to a contestant with the highest type. Each contestant can take a costly test that would reveal his type.

Chapter 5 demonstrates that, as the number of contestants increases, the probability that some information is communicated rises when the cost of the test is high, and falls when the cost is low. When the number of contestants goes to infinity, the probability of communication approaches a positive constant. Making the test more noisy makes it more likely that contestants take the test. Finally, the chapter also shows that as long as the number of contestants is sufficiently large, it is not optimal for the decision-maker to make the test compulsory.

Chapter 2

Collective Learning with Private Values

2.1 Introduction

The outcomes of many economic reforms and other collective decisions are uncertain when the decision is being made. For example, trade liberalisation can help manufacturing while hurting agriculture, or vice versa - but it is not always evident in advance which sector will gain and which will lose. Reforming higher education system from direct government funding of universities to issuing government-guaranteed study loans can reduce the demand for degrees in engineering and increase the demand for life sciences, or the other way around - but the direction of change may be ex ante uncertain. Allocation of research funding, adoption of environmental regulations, and investment in long-term infrastructure projects are other examples of decisions whose consequences are not known in advance.

The common feature of these scenarios is that individual members of the decision-making body know their payoffs, and those of other members, from the decision under each of the outcomes. But the outcomes themselves are unknown. Yet the decision-making body can vote to *collectively* learn this information. For instance, the decision on an economic reform can be delayed until more information about its effects arrives. A pilot project can be implemented to learn about the effects of the reform. Expert advice may be requested[1]. But when will the group choose to learn, and when will it choose to vote "in

[1] An example of this is a decision of an academic hiring committee on whether to give a candidate a job offer. The committee members' preferences often depend on the skills and interests of the candidate, and individual members may disagree over what constitutes an ideal candidate. It is often impossible for individual members to find out additional information about the candidate, but the committee can vote to request this inofmation from the candidate's department.

ignorance”?

This chapter models the decision of a committee that needs to vote on a reform. If adopted, the reform will give every member a private payoff which depends on a binary state of the world. Individual payoffs in each state are commonly known, but the state is initially unknown. The reform, once adopted, is irreversible, but the committee can decide to delay the vote on the reform. As long as the vote on the reform is being delayed, a binary public signal can arrive at any time and reveal the state. The arrival time of the signal follows an exponential distribution. As long as the signal has not arrived, committee members are updating their common belief about the state. Delaying the decision is costless, and there is no discounting.

Will the committee ever vote not to learn the state? Clearly, if committee members have similar preferences, they will weakly prefer to learn the state before making the decision. However, the following example illustrates a situation in which preferences are heterogeneous and the committee prefers not to learn the state. Suppose the two states $\omega \in \{X, Y\}$ are a priori equally likely, the committee consists of three members, and the decisions are made by simple majority voting. For simplicity, suppose that if the committee decides to wait and learn the state, then the state is immediately revealed. Thus, the committee first votes on whether to learn the state prior to deciding on the reform, and then votes on the reform itself. If the committee rejects the reform, its members receive no payoff. If the reform goes ahead, the payoffs of its members are as follows:

	Payoff in state X	Payoff in state Y
Anna	3	-1
Bob	-1	3
Claire	-3	-3

If the committee votes to learn the state before deciding on the reform, then in every state, the majority of voters are against the proposal. Thus a choice to acquire information gives each agent a zero payoff. If the committee votes without having learned the state, then Anna and Bob each get an expected payoff of 1 if the decision is positive. Hence, when the committee votes on the reform in ignorance, the reform is approved. Therefore, Anna and Bob will receive a strictly higher expected payoff if information is not acquired. The majority thus votes against learning the state, so the decision on the reform is made in ignorance. We can say that the committee has a *collective preference for ignorance*[2].

[2] Note that in this example, learning the state would be an optimal choice from a utilitarian point of view, because the sum of payoffs from the reform is negative in each state. However, it is easy to model a situation in which it would be optimal not to learn the state (e.g. if Claire's payoffs are -1 in each state) and the collective decision remains the same.

The key factor behind this outcome is that the reform is accepted in either state, but is rejected ex ante. Thus, information moves the collective decision away from what is preferred ex ante. As the chapter will show, this result can be generalised: given an arbitrary distribution of state-dependent individual payoffs across voters, the committee will choose to vote on the reform without learning the state when decisions on the reform in the two states are the same[3].

The chapter produces four main results. First, it establishes a simple condition on the distributions of individual payoffs across voters that is necessary and sufficient for the committee to vote not to learn the state. It turns out this happens when preferences of committee members are sufficiently heterogeneous. More precisely, ignorance is a collective decision whenever the committee is more fractionalised on the state-relevant dimension than on the state-irrelevant dimension. This implies that decisions on divisive issues[4] are likely to be made without careful consideration of their consequences. More generally, the chapter suggests that fractionalisation and preference heterogeneity have an important effect on collective decisions. Specifically, in fractionalised societies it is likely that decisions are made in haste, that less expert advice is requested, that reforms are enacted or rejected without researching their potential effects, and that there is less public debate on reforms and laws. In this way, the chapter adds to the large literature that examines the effect of social fractionalisation on various socioeconomic indicators[5][6].

This result is also related to a large literature on information aggregation through voting. A number of studies have looked at whether voting can aggregate information that is dispersed among voters, in the spirit of the Condorcet jury theorem[7]. As Bhattacharya (2013*b*) shows, voting will not in general aggregate information if group members' preferences are heterogeneous. The present chapter proves that if individual preferences are sufficiently heterogeneous (roughly speaking) then, in addition, the group will also vote not to learn payoff-relevant information.

Second, the chapter shows that there exists *a blessing of quick disappointment.* That is, as long as equilibrium payoffs depend on the rates at which the two states are revealed, voters whose least preferred state is revealed faster have higher expected payoffs.

[3]The result remains essentially the same if instead of choosing to learn the binary state, the committee is choosing two acquire an arbitrary binary signal. Furthermore, the results are broadly similar if the state space is arbitrary and the committee is choosing to acquire an arbitrary partition of the state space. These extensions are discussed in Section 8.

[4]Divisive in the sense that there is no general agreement on which outcome (state) is preferred to the other.

[5]Indicators such as quality of governance, economic growth, corruption, risk of civil war, and provision of public goods. See Mauro (1995), Easterly and Levine (1997), Collier (2001), Alesina et al. (2003).

[6]A few studies in political science Anderson and Paskeviciute (2006) and organisational psychology Mannix and Neale (2005) suggest that greater heterogeneity within a society or a group may indeed be associated with lower engagement in public discussion and lower level of information exchange.

[7]See, for example, Austen-Smith and Banks (1996), Feddersen and Pesendorfer (1997), Goertz and Maniquet (2011), Bouton and Castanheira (2012).

Third, the chapter proves that the optimal voting rule is simple majority or supermajority[8]. In a setting without learning, a low voting threshold (possibly below 0.5, i.e. below the simple majority rule) is the (utilitarian) optimal solution when each voter who benefits from the reform receives a high payoff, while each voter who loses from the reform loses only a small amount. In reality, however, voting thresholds less restrictive than simple majority are rarely observed. As this chapter shows, when learning is present, a simple majority or supermajority rule is always optimal. I also show that greater supermajority is optimal when individual gains from the reform are much larger or much smaller than individual losses, or when the ratio of gains to losses is more uncertain.

Fourth, the chapter shows that a rule that makes learning compulsory regardless of the committee's decision is optimal when there exists a minority of voters with a large stake in the collective decision. Hence, a constitutional commitment to transparency can be a mechanism that protects minorities when doing so is socially beneficial from a utilitarian point of view.

A number of papers have previously looked at collective decisions that involve learning. Strulovici (2010) examines the problem facing a committee that, in every period, needs to choose between a safe option and a risky option. Committee members do not initially know their preferences, but they gradually learn them when the risky option is exercised. The decision to exercise the risky option is reversible[9]. Fernandez and Rodrik (1991) use a similar approach in an applied context, examining a collective decision to adopt a risky reform[10]. Unlike these studies, in this chapter exercising the safe option (i.e. delaying the decision) leads to learning, while exercising the risky option (adopting the reform) is irreversible and ends the game. Furthermore, in this chapter, voters know their preferences and those of other voters - only the payoff-relevant state is unknown. This allows me to describe preference distributions under which the committee votes against learning, as well as to compare the values of agents with different preferred states[11].

Lizzeri and Yariv (2013) look at a jury that faces a choice in every period between continuing to gather information and making a decision.They find that greater heterogeneity leads to more information acquisition. Unlike this chapter, they restrict attention to a setting in which jury members have preferences that are monotone in the state - everyone prefers to acquit an innocent defendant and to convict a guilty one. In such a setting, information will always be acquired if it is costless; in Lizzeri and Yariv (2013)

[8]Under certain restrictions on the distribution of payoffs, desribed in Section 6.

[9]Godefroy and Perez-Richet (2013) also develop a model in which individuals do not know their preferences ex ante.

[10]Dewatripont and Roland (1995) also look at adoption of reforms with uncertain outcomes.

[11]Messner and Polborn (2012) similarly consider a choice between delaying the decision (and thus learning some information) and adopting a proposal in the first period. They too, unlike this chapter, consider the case in which individual payoffs are unknown ex ante and independently distributed across voters.

the decision to stop gathering information is driven by the fact that it is costly. Unlike Lizzeri and Yariv (2013), I look at cases in which voters do not necessarily prefer the same decision in each state, and show that when voters are strongly fractionalised based on their preferred state, ignorance can be a collective decision even when learning is costless.

This chapter is also related to a literature on collective search[12] in which a committee can choose to acquire further information by continuing the search for another period. The benefit of having more information can, however, be outweighed by the cost of foregoing the payoff in the current period. In contrast to the collective search approach, in this chapter deciding to learn the state before deciding on the reform does not entail any change in payoffs from the reform. The decision to stay uninformed is driven not by payoffs that are lost when information is acquired, but purely by the effect of that information on the collective decision.

The fact that public information can reduce individual expected payoffs has been noticed by Hirshleifer (1971), who showed that risk-averse individuals may be worse off when when information eliminates insurance opportunities[13]. This chapter also looks at situations in which information has negative value, although the focus here is on the collective decision to learn.

More broadly, this chapter is also related to research on acquisition of private information by individual members of committees[14], as well as to studies of information exchange ampng committee members[15]. Also related is the literature on private information acquisition in coordination games[16].

The rest of the chapter is structured as follows. Section 2.2 introduces the model of collective learning. Section 2.3 analyses the baseline version of that model, in which the two states are revealed at the same rate; it shows that the model can be reduced to a simple one-shot framework and establishes necessary and sufficient conditions for ignorance to be a collective decision. Section 2.4 discusses these conditions and describes the groups and societies that are likely to opt for ignorance. Section 2.5 extends the analysis to the case in which one state is revealed faster than the other. Section 2.6 examines the optimal voting rule. Section 2.7 looks at the effect of a rule that makes learning compulsory for the committee. Section 2.8 provides various extensions to the basic framework. Finally, Section 2.9 concludes.

[12] Albrecht et al. (2010); Compte and Jehiel (2010); Moldovanu and Shi (2013)

[13] See also Gersbach (1991) and Gersbach (2000), who show that in a voting framework, information can make some or all agents worse off.

[14] Persico (2004); Gerardi and Yariv (2008); Gershkov and Szentes (2009); Gersbach and Hahn (2012).

[15] Visser and Swank (2007); Gerardi and Yariv (2007).

[16] Dewan and Myatt (2008); Hellwig and Veldkamp (2009); Myatt and Wallace (2012).

2.2 Model

A group of voters I is considering whether to approve a reform[17]. As long as the reform is not approved, every voter receives zero payoff. When the reform is approved, each voter's payoff depends on the binary state of the world $\omega \in \{X, Y\}$. Each voter $i \in I$ receives a payoff (von Neumann-Morgenstern utility) of x_i in state X and a payoff of y_i in state Y. These payoffs can be positive or negative. Let $x \equiv (x_1, x_2, ...)$ and $y \equiv (y_1, y_2, ...)$ denote vectors of individual state-dependent payoffs. The state is unknown, and I denote by π a common belief that the state is X. Let π_0 be the initial belief. Aside from the state, all aspects of the game (including individual payoffs) are common knowledge.

There is continuous time t; the discount rate is zero. The share of votes necessary to make a positive decision is $\gamma \in [0, 1]$. At every point $t \in [0, +\infty)$, agents can choose whether to put the reform to vote or to delay the decision on the reform. Once they have chosen not to delay it further, the reform is put to vote and is either approved or rejected.

Specifically, at every time t each agent $i \in I$ selects a pair of actions $(\beta_{i,t}\ ,\ a_{i,t}) \in \{0, 1\}^2$. Action $\beta_{i,t}$ is the agent's vote on whether to put the reform to vote at time t or to delay the vote on the reform. Action $\alpha_{i,t}$ is her vote on whether to adopt the reform once the group decides to stop delaying the vote. As long as the share of agents who select action $\beta_{i,t} = 1$ is below γ, the decision is delayed, voters receive no payoff and the game progresses further. Once the share of voters who select $\beta_{i,t} = 1$ reaches γ, learning stops and the decision on the reform is made by counting votes in favour and against the reform. In that case, if the share of voters who chose $\alpha_{i,t} = 1$ is below γ, the reform is rejected, and the game ends with every voter receiving zero payoff. If the share of voters who have selected $\alpha_{i,t} = 1$ (i.e. have voted in favour of the reform) is at least γ, the reform is accepted. In that case the game ends with each $i \in I$ receiving x_i and y_i if the state is X and Y, respectively. Hence, γ describes a voting rule that the group uses to reach decisions - both the decision on whether to delay the vote on the reform, and the decision on whether to approve or reject the reform.

As long as the group is voting to delay the decision, they gradually learn the state. This happens in the following way: at every point, signal $\sigma \in \{X, Y\}$ can arrive and reveal the state. The arrival time of the signal corresponds to jump time of a Poisson process. If $\omega = X$, signal Y never arrives, and signal X arrives with intensity λ_X. If $\omega = Y$, signal Y never arrives, and signal X arrives with intensity $\lambda_Y \leq \lambda_X$. Thus, in state $\omega \in \{X, Y\}$,

[17] For some results, it is more natural to assume that I is a finite set; for others, it is more straightforward to see I as a continuum.

the probability that signal $\sigma = \omega$ arrives during a time interval t is $1 - e^{-\lambda_\omega t}$, and the probability that signal $\sigma \neq \omega$ arrives during that interval is zero.

When the signal arrives, it is observed by all voters. Since a signal can only arrive in the corresponding state, arrival of a signal perfectly reveals the state, shifting the belief π to 0 or 1.

I restrict attention to Markov-perfect equilibria with the common belief π as the state. This implies that at the equilibrium the action of every agent at time t only depends on the probability π that the state is X at time t. Voter i's Markov strategy is thus a pair of functions (β_i, α_i), where $\beta_i : [0,1] \to \{0,1\}$ maps common beliefs that the state is X to a vote on whether to delay the decision; while $\alpha_i : [0,1] \to \{0,1\}$ maps beliefs to a vote on whether to adopt the reform. I will restrict the set of admissible Markov strategies to pairs of functions (β_i, α_i) such that the preimages of 0 and of 1 under β_i and α_i are finite unions of nondegenerate intervals. It will be shown later that at equilibrium, these strategies have a simple cutoff form.

As usual in voting games, this game has many equilibria. In line with much of the literature on such games, I only consider equilibria in which weakly dominated strategies are eliminated. Thus, every agent votes sincerely, as if she were the only one deciding the outcome. Therefore, at a belief π, voter i picks $\alpha_i(\pi) = 1$ if and only if her expected payoff from adopting the proposal at a belief π is positive - i.e. if $\pi x_i + (1-\pi) y_i \geq 0$. At the same time, voter i picks $\beta_i(\pi) = 1$ at a belief π if and only if her equilibrium value from continuing to wait is lower than her equilibrium value of stopping.

Hence the decision to adopt the reform is two-stage. First, the group must decide to stop learning and make vote on the reform. Then they need to vote in favour of the reform. This setup presumes that the group can commit to rejecting a proposal. It is possible to imagine an alternative setup, in which he group cannot make such a commitment[18]. In that case, at each time t each voter chooses one action $\alpha_{i,t} \in \{0,1\}$. The reform is adopted (ending the game and giving each $i \in I$ a payoff of x_i or y_i, depending on the state) when the share of voters who have picked $\alpha_{i,t} = 1$ is at least γ; otherwise, the game continues. I will examine this setup later in the chapter.

2.2.1 Evolution of beliefs

Suppose that $\lambda_X = \lambda_Y$. In this case, both states are revealed at the same rate. Then as long as the signal does not arrive, the belief π that the state is X does not change and stays at π_0. Once a signal arrives at some time t^*, the state is revealed, and π jumps to

[18]This applies, for example, to situations in which a parliament can implement a reform that is difficult to reverse but cannot pass a law that forbids it to ever make such a decision in future.

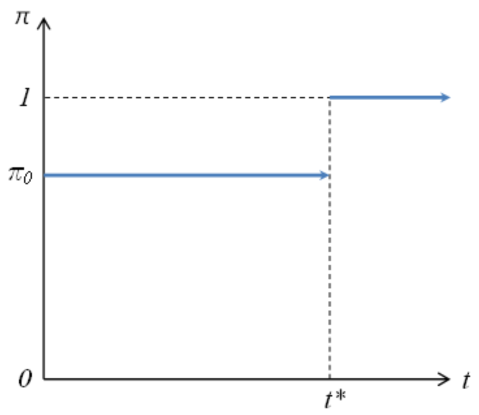

FIGURE 2.1: Evolution of beliefs when $\omega = X$ and $\lambda_X = \lambda_Y$.

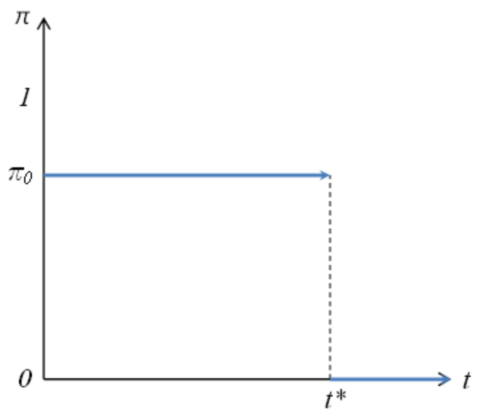

FIGURE 2.2: Evolution of beliefs when $\omega = Y$ and $\lambda_X = \lambda_Y$.

one (if the state is X) or to zero (if the state is Y), and stays there forever. For the case when $\lambda_X = \lambda_Y$, Figures 2.1 and 2.2 show the typical paths of beliefs when the state is, respectively, X and Y.

In this case each voter's Markov strategy reduces to a choice of a pair of actions (b_i, a_i) at beliefs π_0, 0, and 1. Then at any Markov-perfect equilibrium, the group has two options. First, it can decide to vote on the reform immediately at $t = 0$, and either approve or reject the reform without learning the state. Alternatively, it can wait until a signal arrives and reveals the state; at that point the reform is either approved immediately, or never approved[19]. At no other point will the group decide to stop. Since waiting is costless and there is no discounting, this setting is thus equivalent to a one-shot model in which the group has a choice of acquiring a public signal about the state of the world before making the decision.

Suppose now that $\lambda_X > \lambda_Y$. Since the state X is revealed faster, the belief π is falling as group is waiting. The evolution of belief is determined by Bayes' law, under which the probability that the state is X conditional on no signal arriving by the time t equals

$$\pi_t = \frac{\pi_0 \mathrm{e}^{-\lambda_X t}}{\pi_0 \mathrm{e}^{-\lambda_X t} + (1 - \pi_0)\,\mathrm{e}^{-\lambda_Y t}} = \frac{\pi_0}{\pi_0 + (1 - \pi_0)\,\mathrm{e}^{(\lambda_X - \lambda_Y)t}} < \pi_0$$

If the group is continuing to delay the decision, the belief evolves along this path until at some time t^* a signal arrives and reveals the state. At that point the belief jumps to one or zero and stays there forever. The typical paths of beliefs when the state is X and Y are shown in Figures 2.3 and 2.4, respectively.

[19]In this case, never approving the reform can mean either calling a vote and rejecting the proposal; or delaying the decision forever. These two outcomes deliver the same payoffs to all voters, and hence I do not distinguish between them.

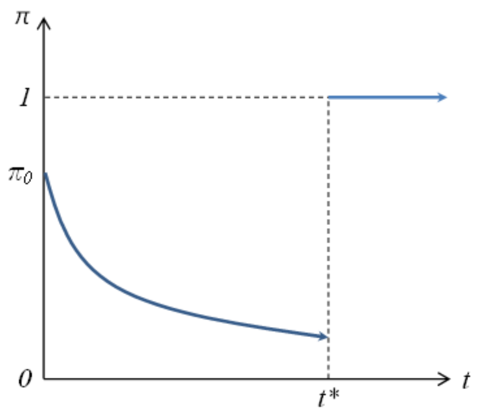

FIGURE 2.3: Evolution of beliefs when $\omega = X$ and $\lambda_X > \lambda_Y$.

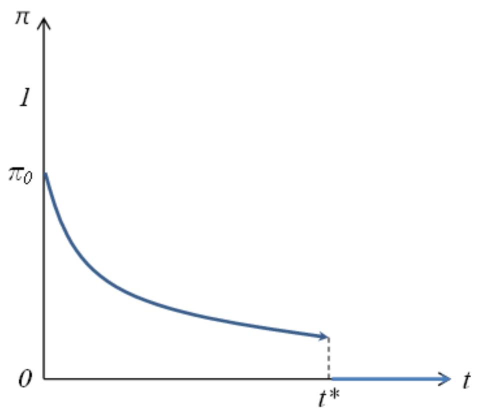

FIGURE 2.4: Evolution of beliefs when $\omega = Y$ and $\lambda_X > \lambda_Y$.

As in the previous case, when the state is revealed the game essentially ends - at a Markov-perfect equilibrium, the reform is either adopted immediately, or never adopted. As long as the signal has not arrived, each voter's actions $\beta_i(\pi)$ and $\alpha_i(\pi)$ can change with π - and by extension, the collective decision also changes.

For now, I will focus most of the analysis on the case when the rates of learning are the same for both states. I will later extend the model to the case when the rates of learning are asymmetric, i.e. when $\lambda_X > \lambda_Y$.

2.3 Symmetric Rates of Learning

2.3.1 Preference for Ignorance

Let $g_\gamma : \mathbb{R}^I \to \{0,1\}$ denote a function that maps the vector of expected payoffs from adopting the reform at a given belief to the collective decision on the reform, given the voting rule γ. As each agent votes sincerely, she will vote in favour of adopting the reform if and only if her expected payoff from the reform is positive, given the belief π. Thus, if the group has decided to stop learning and make a vote on the reform at a belief π, the reform is adopted if $g_\gamma[\pi x + (1-\pi) y] = 1$, and is rejected if $g_\gamma[\pi x + (1-\pi) y] = 0$. Thus, for a given vector of expected payoffs, the function $g_\gamma(\cdot)$ takes a value of 1 if and only if the share of positive elements in that vector is at least γ. For $\gamma = \frac{1}{2}$ (simple majority rule) $g_\gamma(\cdot)$ indicates whether the median of a vector is positive. For other values of γ, $g_\gamma(\cdot)$ indicates whether the corresponding quantile of the vector of payoffs is positive.

Note that $g_\gamma(\cdot)$ has the following properties, which will be useful in the subsequent analysis:

1. For any $\gamma \in [0,1]$, any scalar $\lambda > 0$, and any payoff vector $z \in \mathbb{R}^I$, it holds that $g_\gamma(z) = g_\gamma(\lambda z)$.

2. When $\gamma = \frac{1}{2}$ and I is finite and odd-sized, then for any payoff vector $z \in \mathbb{R}^I$ that does not contain zeroes, it holds that $g_\gamma(-z) = 1 - g_\gamma(z)$.

The first property says that the collective decision is invariant to rescaling of payoffs. The second says that for a simple majority rule, if the reform is adopted under a given vector of expected payoffs, then it is rejected under a vector of opposite expected payoffs.

As explained previously, when the two states are revealed at the same rate, the decision problem of the group can be reduced to a choice between learning the state and not learning it, followed by a choice between adopting and rejecting the reform. If the group has decided to learn the state and the state is X, then the reform is approved if and only if $g_\gamma(x) = 1$. Similarly, if the group decides to wait and the state turns out to be Y, then the reform goes ahead if and only if $g_\gamma(y) = 1$. Thus, if the collective decision is to learn the state, the ex ante expected payoff of agent i equals

$$\pi_0 x_i g_\gamma(x) + (1-\pi_0)\, y_i g_\gamma(y)$$

If the group decides not to learn the state, then the reform is approved if and only if $g_\gamma(\pi_0 x + [1-\pi_0]\, y) = 1$. In that case, agent i's expected payoff equals

$$(\pi_0 x_i + [1-\pi_0]\, y_i)\, g_\gamma(\pi_0 x + [1-\pi_0]\, y)$$

Sincere voting implies that each $i \in I$ supports putting the reform to vote at time 0 (before the state is revealed) if and only if learning the state decreases her expected payoff. The value of ignorance for agent i - i.e th gain in i's expected payoff from putting the reform to vote at $t = 0$ instead of delaying the decision - then equals

$$(\pi_0 x_i + [1-\pi_0]\, y_i)\, g_\gamma(\pi_0 x + [1-\pi_0]\, y) - \pi_0 x_i g_\gamma(x) - (1-\pi_0)\, y_i g_\gamma(y) \equiv d_i$$

Voter i votes in favour of deciding on the proposal at $t = 0$ if $d_i > 0$, and votes against it if $d_i < 0$. Then the committee collectively chooses to delay the decision and learn the state whenever $g_\gamma(d) = 0$, where $d \equiv (d_i)_{i \in I}$ is the vector of net gains from ignorance for all agents. When $g_\gamma(d) = 1$, the committee collectively chooses to vote on the reform at time 0, without learning the state - in other words, the group has a *collective preference for ignorance.*

Before proceeding further, we need to consider the possibility that voters are indifferent between some of the decisions. In particular, when $g_\gamma(x) = g_\gamma(y) = g_\gamma(\pi_0 x + [1-\pi_0]\, y)$

- i.e. when the group votes on the reform the same way in either state and also ex ante - then $d_i = 0$ for all $i \in I$. In that case all voters are indifferent between learning and not learning the state. Therefore, the subsequent analysis will distinguish between a weak and a strict collective preference for ignorance.

With this distinction in mind, we can derive the necessary and sufficient conditions for the group to decide not to learn the state.

Proposition 1. *Suppose that $\gamma = \frac{1}{2}$ and I is finite and odd-sized. Then for any $\pi_0 \in (0,1)$, and any λ_X, λ_Y such that $\lambda_X = \lambda_Y$, the committee has a weak preference for ignorance if and only if $g_\gamma(x) = g_\gamma(y)$. Furthermore, the committee has a strict preference for ignorance if and only if $g_\gamma(x) = g_\gamma(y) \neq g_\gamma(\pi_0 x + [1-\pi_0]y)$.*

Proof. See Appendix.

In words, the committee weakly prefers making a decision without information whenever the decisions under the two realisations of the signal are the same. The committee strictly prefers not acquiring information whenever the decisions under the two realisations of the signal are the same and also different from a decision that is made in ignorance.

Intuitively, the group wants to wait and learn the state when knowing the state affects the decision on the reform. If the decision is the same in both states, two cases are possible. First, that decision can also be the same as the decision on the reform in ignorance - in this case, information has no effect on the outcome, and the committee weakly prefers not to have it. Second, the decisions in both states can be be different from the decision that is preferred by the majority ex ante. In this case, information moves the collective decision on the reform away from the decision was optimal ex ante - thus, the majority will prefer not to have that information.

Corrollary 1 (). *Under a simple majority rule, whenever there is a strict preference for ignorance, the preferred decision of the median voter in either state will never be implemented.*

Proof. Strict preference for ignorance exists when the decision on the reform preferred by the median voter in either state is different from the ex ante decision. Under a strict preference for ignorance, the ex ante decision on the reform will be the one that the group will implement.

This result is, of course, different from what the median voter theorem suggests. When the group can endogenously determine whether they want to acquire information about

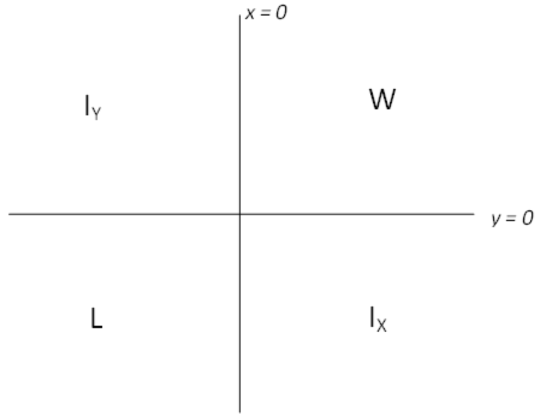

Figure 2.5: Distribution of preferences.

the state, and the median voter strictly prefers not to acquire that information, then agents who are median voters in each state will *never* have their preferred alternative chosen.

2.3.2 Preference Distributions

The analysis above has looked at how decisions upon receiving different realisations of the signal affect the willingness of the committee to acquire that signal. The primitives of the model, however, are not these decisions, but individual preferences. We can now examine the question of what distributions of preferences among voters give rise to a collective preference for ignorance.

Suppose that the voting rule is simple majority (i.e. $\gamma = \frac{1}{2}$), we can look at the actual distributions of individual preferences that induce a collective preference for ignorance. Recall that the preferences of any agent i are described by a pair (x_i, y_i) of his payoffs from accepting the proposal under the two states. The distribution of preferences over the group is then the distribution of agents over the (x, y) space. Figure 2.5 shows the (x, y) space of state-dependent payoffs.

In Figure 2.5, letters W, L, I_Y, and I_X indicate the sets of agents whose preference points are in areas bounded by thick lines. Thus, W represents the set of sure winners - they have a positive expected utility from adopting the proposal after any signal. L represents the set of sure losers, who prefer the proposal to be rejected after any signal. We can refer to the sets W and L as the sets of *committed* voters, or *partisans*. I_X and I_Y are the sets of *independent* voters, whose preferred decision changes depending on the signal. I_X are independent voters that prefer the proposal to be accepted when the signal is 1 (i.e. when the state is more likely to be X), but not when the signal is 0. I_Y

are independent voters who get a positive expected utility from the proposal when the signal is 0, i.e. when the state is likely to be Y.

Suppose for simplicity that the mass of those for whom $px + (1-p)\,y = 0$ or $(1-p)\,x + py = 0$ is zero, i.e. that (almost) nobody is indifferent when either of the signals is received. For a given set of voters S, let $\#S$ denote the share of voters who belong to the set S. Then the following result holds:

Proposition 2. *Under the simple majority rule, the committee has a weak preference for ignorance if and only if* $|\#I_X - \#I_Y| \leq |\#W - \#L|$

Proof. See Appendix

This is a necessary and sufficient condition on individual payoff distribution to induce a weak preference for ignorance. Appendix 1 provides a necessary and sufficient condition for a strict preference for ignorance to exist.

Thus, information will not be acquired if and only if the difference between the numbers of independents of the two types is smaller than the difference between the number of sure winners and the number of sure losers. Hence, ignorance will be the collective decision when the distribution of individuals is "relatively symmetric" along the "northwest-southeast" direction (i.e. when I_X and I_Y are similar), and "relatively asymmetric" along the "northeast-southwest" direction (when W and L are different).

2.3.3 The Case Without Commitment

The preceding analysis looked at a situation in which the group could cease learning and then reject the reform irreversibly. In many situations, however, a reform can be irreversibly adopted without learning information about its consequences; yet it may be impossible to reject the reform and ensure that it will never be reconsidered in future, when the collective belief changes. For example, a parliament of a country can vote to join a currency union or can delay the decision on joining, but it may be impossible to pass a law saying that the country will never join.

If it is impossible to commit to rejecting the reform, then at any time t, each voter chooses whether she would like to adopt the reform or delay the decision and continue learning. Hence, the committee can never decisively reject the reform. But it is possible that the committee chooses to delay the decision forever, in which case the reform is never adopted. In terms of the model described above, inability to commit to rejecting the reform means that each voter i only chooses $\alpha_{i,t} \in \{0,1\}$ at every time t.[20] Once the share of voters who choose $\alpha_{i,t} = 1$ reaches γ, the reform is adopted and the game ends. Until then, the committee continues learning. Voter i's Markov strategy is then a function $\alpha_i : [0,1] \to \{0,1\}$.

When $\lambda_X = \lambda_Y$, the committee, by the same logic as in the case with commitment, has only two alternatives: stop learning and adopt the reform immediately, or delay the decision until the state is revealed (after which the reform is either adopted immediately or delayed forever). Then the expected payoff of agent i if the committee votes to adopt the reform at time 0 is

$$\pi_0 x_i + (1 - \pi_0)\, y_i$$

and if they decide to delay the vote, i's payoff is

$$\pi_0 x_i g_\gamma(x) + (1 - \pi_0)\, y_i g_\gamma(y)$$

These are equivalent when $g_\gamma(x) = g_\gamma(y) = 1$, making all agents indifferent. If $g_\gamma(x) = 1$ and $g_\gamma(y) = 0$, then i benefits from adopting the reform early if and only if $y_i > 0$. But if $g_\gamma(y) = 0$, the share of such voters is lower than γ - hence, the committee votes to wait for the state to be revealed, and then adopts the reform if and only if the state turns out to be X. Similarly, the committee decides to learn when $g_\gamma(x) = 0$ and $g_\gamma(y) = 1$. When $g_\gamma(x) = g_\gamma(y) = 0$, voter i's gain in payoff from adopting the reform in ignorance is $\pi_0 x_i + (1 - \pi_0)\, y_i$. Then the committee approves the reform at time 0 whenever $g_\gamma(\pi_0 x + [1 - \pi_0]\, y) = 1$.

This result is summarised in the following proposition:

Proposition 3. *Suppose the committee cannot commit to reject the reform. Then for any* $\gamma \in [0,1]$*, any* $\pi_0 \in (0,1)$*, and any* λ_X, λ_Y *such that* $\lambda_X = \lambda_Y$*, the committee has a strict preference for ignorance if and only if* $g_\gamma(x) = g_\gamma(y) = 0$ *and* $g_\gamma(\pi_0 x + [1 - \pi_0]\, y) = 1$.

Proof. See Appendix.

In words, the reform is approved without learning the state whenever it is rejected in each state of the world while being preferred ex ante. This result holds for any voting rule γ.

[20] Equivalently, we can say that inability to commit restricts $\alpha_{i,t}$ and $\beta_{i,t}$ in such a way that $\alpha_{i,t} = \beta_{i,t}$ for every pair (i,t).

It is easy to see that this result is similar to the result in Proposition 1, except that without commitment power, making a decision on the reform in ignorance is only possible when that decision is positive.

2.4 Discussion of Results

Consider again the result in Proposition 2.5. What does it mean for a group to have a distribution of preferences under which the numbers of independent voters are more similar than the numbers of partisans? One way of interpreting this condition is to note that this outcome is likely when the states X and Y are possible outcome of the reform, and committee members do not share a common opinion on which outcome is better. This happens when individual utilities from the reform are not monotone in the state. On the other hand, if all voters receive a higher payoff from the reform in state X than in state Y,[21] then I_Y is empty, and the committee can at most weakly prefer ignorance[22].

Thus, decisions on divisive issues - divisive in the sense that there is much disagreement among voters over what outcome of the decision is preferred - are likely to be made in haste and without careful consideration of their consequences. For example, suppose that a healthcare reform is being discussed in the parliament, and it is unclear whether it will make abortion more or less accessible. If the parliament is largely split between supporters and opponents of abortion, the reform is likely to be adopted or rejected in haste.

Another way to interpret Proposition 2 is to refer to the index of social fractionalisation[23], widely used in development literature. For a society divided into different groups, the index of fractionalisation measures the probability that two randomly selected individuals belong to different groups. When the society consists of two groups, this index is higher when the sizes of the two groups are more similar. Proposition 2 then says that ignorance will be a collective decision if fractionalisation on the state-relevant dimension is larger than fractionalisation on the state-irrelevant dimension.

In fact, there is substantial research in development literature on the impact of social fractionalisation on economic growth, corruption, quality of governance, public good provision, and risk of civil war[24]. This chapter suggests an alternative mechanism through

[21] This happens, for example, when each voter's payoff from the reform is a sum of her private value A_i and state-dependent common value $B * \mathbf{1}(\omega = X)$, where $\mathbf{1}(\cdot)$ is the indicator function and $B > 0$.

[22] This is because in this case, ignorance is a collective choice only when $\#I_X \leq |\#W - \#L|$. This happens when $\#W \geq \#L + \#I_X$, or when $\#L \geq \#W + \#I_X$. In these cases the majority of voters either wants to reject the reform or wants to adopt it, regardless of the state. The committee is thus indifferent between learning and not learning the state.

[23] Described in e.g. Montalvo and Reynal-Querol (2005)

[24] See Mauro (1995), Easterly and Levine (1997), Collier (2001), Alesina et al. (2003), and others.

which fractionalisation can affect economic and social outcomes. Specifically, fractionalisation affects the degree to which the society prefers to be informed when making decisions. For a given decision whose outcome depends on the state off the world, a society that is fractionalised on the state-relevant dimension more than on the state-irrelevant dimension will collectively choose to make decisions without learning all the available information - even when learning comes at no cost. For example, societies of that kind are likely to make decisions in haste, and to elicit less advice from experts.

Suppose, for example, that a parliament of a country is considering to join a free trade agreement. There is uncertainty over the effect of free trade on the relative price of agricultural and manufactured goods. If the number of members of parliament supporting the manufacturing sector is similar to the number of MPs supporting agriculture, while the number of ideological opponents of free trade is much larger than the number of committed supporters, then the parliament will not seek information about the likely change of the relative price before voting on the agreement.

Public debate and exchange of information is another way through which a group can learn the state of the world. We may consider a situation in which information about the state is dispersed among voters, with each voter receiving a signal about the state. If individual signals are very imprecise, then all voters have (almost) the same prior belief. By making these signals public, the society as a whole can become more informed about the state[25]. Certain norms and institutions - such as freedom of speech and a strong tradition of public discussion - can make the exchange of signals easier. Proposition 2 suggests that societies that largely agree that some outcomes are better than others are more likely to support the existence of such institutions. On the other hand, societies that are more heterogeneous in terms of their preferences are, ceteris paribus, less likely to collectively support them.

Finally, note that the necessary and sufficient condition on the collective preference for ignorance given in Proposition 2 is similar to the condition on the failure of information aggregation through voting Bhattacharya (2013*b*). Condorcet Jury Theorem states that voting effectively aggregates information that is dispersed among a large number of voters. A number of papers (e.g. Feddersen and Pesendorfer (1997)) have shown that this holds in settings in which individual payoffs from an alternative are monotone in the state - i.e. when all voters agree that an alternative is better in one state than in the other. However, Bhattacharya (2013*b*) shows that when individual preferences are not monotone, information aggregation in general does not hold. This analysis suggests that

[25]One can argue, based on Condorcet Jury Theorem, that voting will aggregate all this information even if voters do not exchange it. But as discussed below, in settings in which a collective preference for ignorance emerges, Condorcet Jury Theorem is not likely to hold.

in these settings, not only does voting fail to aggregate information, but the committee also *chooses* not to learn information when it has an option of doing so.

2.5 Asymmetric Rates of Learning

The previous section has looked at a model of collective learning in which the two states were revealed at the same rate. Often, however, this is not the case. Consider, for example, a funding body that is supporting two research projects related to power production. One project is focused on developing a new solar power plant design, while another is researching fusion power. The funding committee would like to focus its resources on the project that is more likely to be successful, but some members are privately interested in supporting fusion power research, while others - in supporting solar power. It is unclear at the moment which project is more likely to succeed, but the decision on funding can be delayed. While the committee is delaying the decision, each project can reach a breakthrough. However, solar power is a near-term project, and if it is indeed successful, a breakthrough is likely to come early. Fusion power research, on the other hand, is a long-term project which can last for a long time without a major breakthrough - even if, in fact, it will eventually be key to producing cheap and abundant power. How long will the committee deliberate before making a decision, and will the supporters of solar power or the supporters of fusion power benefit from asymmetric revelation of information?

In this section I will analyse the group's decision when $\lambda_X > \lambda_Y > 0$ - hence, state X is revealed faster than state Y. Beliefs thus evolve as Figures 2.3 and 2.4 show.

Throughout this section I will assume that the group is unable to commit to rejecting a proposal. Hence, at each belief π an agent can vote either to stop learning and adopt the reform, or to continue learning. Adopting the reform gives each agent $i \in I$ an expected payoff of $s_i(\pi) = \pi x_i + (1-\pi) y_i$.

As before, once the state is revealed, the game effectively ends - the reform is either immediately adopted or never adopted. Before the state is revealed, the collective decision at any belief π will depend on what happens when the state is revealed, given the voting rule γ.

Any strategy profile and voting rule γ must imply a collective decision at a point at which the state is revealed. This decision is described by $g_\gamma(x)$ and $g_\gamma(y)$, which take the value of 1 if the reform is adopted immediately when the corresponding state is revealed, and a value of 0 - if it is never adopted. This describes what happens when the belief π equals, respectively, one or zero. When $\pi \in (0,1)$, any strategy profile implies, for any belief π, a certain value $T(\pi) \in [0, +\infty)$ such that the reform is adopted T units of time later if no signal arrives before that. $T(\pi) = 0$ means that the reform is adopted at the belief π; $T(\pi) = \infty$ means that the committee waits until the state is revealed, at which point its decision is governed by $g_\gamma(x)$ and $g_\gamma(y)$.

We can look at three cases: the one in which the reform is adopted after either state is revealed; the one in which it is adopted in one state but not in the other; and the one in which it is never adopted once either state is known.

2.5.1 Symmetric Positive Decisions

Fix a voting rule $\gamma \in [0, 1]$, and suppose that, given γ and the distribution of individual payoffs within I, the reform is adopted once either state is revealed - i.e. $g_\gamma(x) = g_\gamma(y) = 1$. Then the reform will eventually be adopted regardless of how long the committee chooses to wait. In this case, at any belief all agents are indifferent between voting for or against the reform.

2.5.2 Asymmetric Decisions.

Suppose that $g_\gamma(x) \neq g_\gamma(y)$. Then we can show that the committee will wait until the state is fully revealed, as the following proposition shows:

Proposition 4. *Take any committee I, any voting rule γ, and any λ_X, λ_Y such that $\lambda_X > \lambda_Y > 0$, and suppose that the committee cannot commit to reject the reform. If $g_\gamma(x) \neq g_\gamma(y)$, then the committee will delay the decision until the state is revealed.*

Proof. See Appendix.

This result essentially repeats the result in Proposition 3 for the case when $\lambda_X > \lambda_Y$.

2.5.3 Symmetric Negative Decisions.

Suppose that, given γ, the reform is rejected when either state is revealed - i.e. that $g_\gamma(x) = g_\gamma(y) = 0$. Then delaying the decision means that, with some probability, a signal arrives and the reform is rejected.

Any agent for whom $x_i > 0$ and $y_i > 0$ wants to adopt the reform at any belief. Hence, she always votes to stop learning and adopt the proposal - she gains nothing from learning the state, but waiting carries a risk of signal arriving and the proposal being rejected.

Similarly, a voter for whom $x_i < 0$ and $y_i < 0$ always votes to continue learning and against the reform.

Now let us look at voters for whom x_i and y_i have different signs. These are voters who belong to groups I_X and I_Y, i.e. the independent voters. For each of them, we can denote a belief at which voter i is indifferent between approving and rejecting the reform. This belief equals

$$\frac{y_i}{y_i - x_i} \equiv \pi_i^M$$

and is known as the myopic threshold in experimentation models[26].

If the belief were static, voter $i \in I_X$ would have been in favour of adopting the reform if and only if $\pi \geq \pi_i^M$, while voter $i \in I_Y$ would have supported adopting the reform if and only if $\pi \leq \pi_i^M$. Here, however, the belief π is not static. Therefore, a voter who would have otherwise voted to adopt the reform may want to wait to make sure that the state is the one in which her payoff from the proposal is positive. On the other hand, waiting too long means that the state is likely to be revealed, in which case the reform will be rejected. This tradeoff determines, for every independent voter i, the optimal cutoff belief at which she switches her vote from accepting the proposal immediately to delaying the decision.

Consider an independent voter whose preferred state is X. At any belief $\pi < \pi_i^M$, his payoff from adopting the reform is negative. Thus, at any belief below π_i^M, she votes against adopting the reform, hoping that the committee will end up delaying the decision until the state is revealed, in which case the reform never goes ahead. Now suppose that $\pi > \pi_i^M$. In that case, if the committee stops learning and adopts the reform, voter i will receive a positive payoff. If instead they choose to wait, the reform will either be rejected (if a signal arrives), or the group will eventually choose to stop and adopt the reform at some other belief π'. But since the belief is evolving downward, $\pi' < \pi$. Thus, waiting means that the committee either rejects the reform, or adopts it at a belief that is less favourable for agent i. Either of these alternatives is worse for voter i than adopting the reform at the belief π. We can conclude that any $i \in I_X$ votes in favour of the reform if and only if the belief is better for him than his myopic cutoff. This reasoning is summarised in the following proposition:

Proposition 5. *Take any committee I, any voting rule γ, and any λ_X, λ_Y such that $\lambda_X > \lambda_Y > 0$, and suppose that the committee cannot commit to reject the reform. If $g_\gamma(x) = g_\gamma(y) = 0$, then any agent i such that $x_i > 0 > y_i$ votes in favour of the reform if $\pi > \pi_i^M$ and against the reform if $\pi < \pi_i^M$.*

[26] Keller et al. (2005).

Proof. See Appendix.

Now consider a voter $i \in I_Y$. For this voter, stopping at any belief $\pi > \pi_i^M$ gives a negative payoff. So, she prefers to continue waiting. At a belief $\pi < \pi_i^M$, i's payoff from stopping and adopting the reform is positive. However, for an independent voter whose preferred state is Y, there is a benefit in waiting. This is because waiting for some time makes it possible to adopt the reform at a belief that is more favourable for the voter. Unlike an agent whose preferred state is X, voter $i \in I_Y$ benefits from acquiring information. Thus, in general, she will want to continue to delay adoption of the reform for some beliefs below π_i^M. This result is expressed in the following proposition:

Proposition 6. *Take any committee I, any voting rule γ, and any λ_X, λ_Y such that $\lambda_X > \lambda_Y > 0$, and suppose that the committee cannot commit to reject the reform. If $g_\gamma(x) = g_\gamma(y) = 0$, then any agent i such that $y_i > 0 > x_i$ prefers the reform to be adopted if and only if*

$$\pi < \frac{y_i}{y_i - \frac{\lambda_X}{\lambda_Y} x_i} \equiv \pi_i^* < \pi_i^M$$

Proof. See Appendix.

This result describes the belief at which an independent agent whose preferred state is Y would like the reform to be adopted. But depending on the distribution of individual preferences agent $i \in I_Y$ may not be able to force the reform to be adopted at his preferred cutoff π_i^*. This is because even with i's vote, the number of agents who vote in favour of adopting the reform at the belief π_i^* may not be sufficient for the reform to be adopted. On the other hand, i's expected utility from adopting the reform is positive at any belief $\pi \in \left(\pi_i^*, \pi_i^M\right)$. Thus, when faced with the choice between adopting the reform at some belief $\pi \in \left(\pi_i^*, \pi_i^M\right)$ and having the reform never adopted, i may choose to vote to stop information acquisition before the belief reaches π_i^*.

The following example illustrates this idea.

2.5.4 Example

Consider the example in which the committee consists of three voters, Anna, Bob, and Claire, whose preferences are given in the table in the Introduction. Suppose that $\lambda_X > \lambda_Y > 0$, and let $\gamma = \frac{1}{2}$. Then, given the agents' preferences, $g_\gamma(x) = g_\gamma(y) = 0$, i.e. the reform is never adopted once either state is revealed.

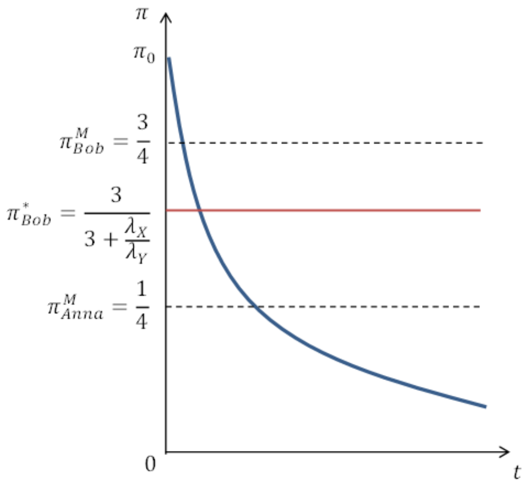

FIGURE 2.6: Collective learning when $1 < \frac{\lambda_X}{\lambda_Y} < 9$.

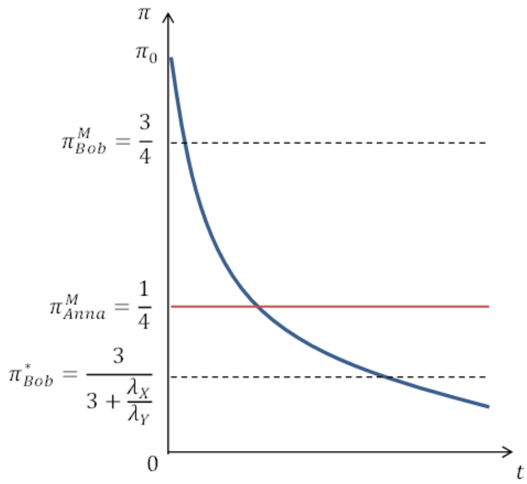

FIGURE 2.7: Collective learning when $\frac{\lambda_X}{\lambda_Y} > 9$.

In this example, Claire always loses from the reform, so she always votes against it. The reform is the adopted whenever both Anna and Bob agree to adopt it. Anna receives a positive payoff in state X and a negative payoff in state Y. She votes to adopt the reform if and only if $\pi > \pi^M_{Anna} = \frac{-1}{-1-3} = \frac{1}{4}$. Bob, on the other hand, receives a positive expected payoff from the reform if $\pi < \pi^M_{Bob} = \frac{3}{3+1} = \frac{3}{4}$. But as long as the state is not revealed, the probability π that the state is X is decreasing - i.e. evolving in a direction that is favourable for Bob. If Bob was the only agent making the decision, he would choose to wait until the belief falls to

$$\pi^*_{Bob} = \frac{3}{3+\frac{\lambda_X}{\lambda_Y}}$$

at which point he would choose to adopt the reform.

Consider first the case when $1 < \frac{\lambda_X}{\lambda_Y} < 9$, which is illustrated in Figure 2.6. Let the initial belief π_0 be above $\frac{3}{4}$. At that belief, only Anna is in favour of adopting the reform, so the reform is not adopted and the committee continues waiting and learning. Gradually, the belief is decreasing. When it reaches $\pi^M_{Bob} = \frac{3}{4}$, Bob's expected payoff from adopting the reform becomes positive. However, at that belief Bob can still gain from learning, and hence he continues to vote against adopting the proposal. The belief then decreases until it reaches π^*_{Bob}. At that point, Bob switches his vote and supporting the reform. Since Anna is still voting in favour of the reform at that belief, the reform is adopted at the belief $\pi^*_{Bob} = \frac{3}{3+\frac{\lambda_X}{\lambda_Y}}$ - unless, of course, a signal arrives and reveals the state prior to that, in which case the reform is never adopted.

Now consider the case in which $\frac{\lambda_X}{\lambda_Y} > 9$, depicted in Figure 2.7. As before, only Anna supports the reform at π_0. As the committee is delaying the vote, the belief eventually falls below π^M_{Bob}. Again, Bob's expected payoff from the reform now becomes positive, but he continues to vote against the reform in order to learn more information. Eventually, the belief reaches $\pi^M_{Anna} = \frac{1}{4}$, which in this case is larger than π^*_{Bob}. At that point, if the belief falls further, Anna would change her vote and start voting against the reform. Bob thus knows that if he does not vote in favour of the reform when $\pi = \pi^M_{Anna}$, then the reform will never be adopted and he will receive zero payoff. Since Bob's expected payoff from adopting the reform at $\pi = \pi^M_{Anna}$ is positive, Bob chooses to vote in favour of the reform at that belief. Thus, Anna and Bob both vote for the reform, and the reform is adopted when $\pi = \pi^M_{Anna}$.

2.5.5 The Blessing of Quick Disappointment

Consider the example above. When $\frac{\lambda_X}{\lambda_Y} > 9$, there are two possible outcomes. First, the state may be revealed before the belief reaches π^M_{Anna}. In that case, the reform is never accepted, and all agents receive zero payoffs. Second, the state may not be revealed by that time, in which case the reform is adopted at a belief π^M_{Anna} - i.e. at the belief at which Anna's expected payoff from the reform is zero. Hence, in expectation, Anna receives a zero payoff. Bob, on the other hand, receives a positive expected payoff, since there is a positive probability of the reform being accepted at the belief $\pi^M_{Anna} < \pi^M_{Bob}$. We can thus conclude that in this simple setting, the agent whose preferred state is revealed faster has a zero expected payoff if the difference in learning rates across the two states is sufficiently large. This easily generalises to a setting in which all independent agents $i \in I_X$ (such as Anna) are identical and all independent agents $i \in I_Y$ are identical too.

More generally, we can see that Anna's decision is predetermined - she votes in favour of the reform if and only if the reform gives her a positive expected payoff. Bob, on the other hand, chooses the point at which he switches his vote from opposing to supporting the reform. Hence, Bob - the voter whose preferred state is revealed slower - is the one who effectively decides when to stop learning and adopt the reform. Bob thus has an advantage.

The result that an agent is better off if his favourite state is revealed slower is actually fairly general, and can be formalised as follows. Take any voting rule γ, and any λ_X, λ_Y. For an agent whose payoff from the proposal in states X and Y are x_i and y_i, respectively, let $u(x_i, y_i, \pi)$ be the equilibrium value function at a belief π. We are interested in how that value depends on the rates λ_X, λ_Y at which states are revealed, and on whether i's preferred state is revealed faster or slower. Clearly, $u(x_i, y_i, \pi)$ depends on the i's preferences (x_i, y_i) - for example, increasing the payoffs under both states increases the expected payoff when λ_X, λ_Y are held constant - so we need to look at comparable cases.

Specifically, we can compare two agents with symmetric preferences - namely agents, one of whom receives a in state X and b in stat Y, while the other receives b in state X and a in state Y, where $a < b$

It turns out that one of the following must hold. Either voters' values do not depend on which state is revealed faster; or if they do, then out of two voters with symmetric preferences, the voter whose least preferred state is revealed faster is better off.

Proposition 7. ***(The blessing of quick disappointment)*** *Take a committee I and suppose it includes voters with preferences $(x_0, y_0) = (a, b)$ and $(x_1, y_1) = (b, a)$, where $a < b$. Then for any $\gamma \in [0, 1]$, either*

(i) for all $i \in I$, $u(x_i, y_i, \pi)$ does not depend on the size of λ_X or λ_Y; or

(ii) $u(a, b, \pi) \geq u(b, a, \pi)$ for all $\pi \leq \frac{1}{2}$ and all λ_X, λ_Y such that $\lambda_X \geq \lambda_Y$.

Proof. If $g_\gamma(x) = g_\gamma(y) = 1$, the reform is always accepted, and each $i \in I$ receives a payoff x_i or y_i, regardless of λ_X or λ_Y. If $g_\gamma(x) \neq g_\gamma(y)$, then by Proposition 4 the committee waits until the the state is revealed, and then, depending on the state, either accepts or rejects the reform. In this case, payoffs again do not depend on λ_X or λ_Y. If $g_\gamma(x) = g_\gamma(y) = 0$, the voter whose preferred state is received slower is better off; the proof of this is given in the Appendix.

2.5.6 Non-Monotone Effect of Voting Rules

What is the effect of changing the voting rule γ on a collective decision? On the face of it, reducing γ makes it is easier for the group to adopt the reform. This should lead to earlier adoption of the reform, with less learning.

Suppose, however that at the current level of γ, $g(x) = g(y) = 0$. Now let us reduce γ, making the decision rule less conservative. Initially, this expands the set of beliefs at which the reform is adopted. Hence, in general, the amount of learning decreases, and the proposal is adopted faster - we can say that the decision becomes less conservative. At some point, however, γ may become sufficiently low that either $g(x)$ or $g(y)$ becomes 1. If that happens, then by Proposition 4 the committee chooses instead to wait until the state is revealed, thus opting for maximum learning. Hence, the even though the decision-making criterion has become less stringent, the reform is adopted later and with greater consideration.

Thus, making the decision rule less conservative can lead to a more conservative decision.

2.6 Optimal Voting Rule

This section will examine the choice of a voting rule that would maximise the utilitarian welfare of the voters.

Suppose that $\lambda_X = \lambda_Y$, and the committee cannot commit to rejecting the reform. Let us simplify the payoff structure in the following way: suppose that every voter belongs to either group I_X or group I_Y. Suppose further that $\pi_0 x_i + (1-\pi_0)\, y_i$ - the ex ante expected payoff from the reform - is the same for all $i \in I$.[27] Thus, voters agree about the value of the reform ex ante; but ex post, every voter gains from the reform in one state and loses in the other, and ex post payoffs can vary across agents. This is not unlike the setting in Fernandez and Rodrik (1991).

Consider the problem of a social planner that needs to select voting rule γ to maximise the utilitarian expected welfare. The social planner does not know the state, nor does he know individual payoffs - they are drawn from some distribution subject to the restrictions described above. When agents vote, however, they do know the payoffs x_i and y_i of all the committee members. Thus, once the voting rule is chosen, the nature draws the state of the world and individual payoffs, and these payoffs become common knowledge. Afterwards, the game proceeds as described earlier in the chapter. Let $z \equiv \pi_0 x_i + (1-\pi_0)\, y_i$ be the realised (common) expected payoff from the reform, and let $s \in [0,1]$ be the realised share of agents who belong to group I_X. Assume that s and z are distributed independently[28], and that the distribution of s has full support on the $[0,1]$ interval. To make the last assumption meaningful, I will assume in this section that the set of voters I is a continuum with a total mass of one. Let $\mu_X \equiv \int\limits_{i\in I} x_i di$ and $\mu_Y \equiv \int\limits_{i\in I} y_i di$ denote the overall utilitarian welfare from adopting the reform in state X and Y, respectively, given the realised payoffs.

This corresponds to a situation in which the committee needs to make a number of decisions in the future, and each decision is associated with a different distribution of payoffs. The problem of the social planner can then be interpreted as the problem of finding the optimal constitutional rule.

In a setting without learning, the solution is fairly straightforward. If losses from the reform tend to be concentrated and gains are likely to be dispersed, the optimal rule is supermajority rule, because it would protect the interests of a minority who are set to lose much if the reform goes ahead. Similarly, low γ, possibly lower than 0.5, is optimal when gains tend to be concentrated and losses are dispersed. In reality, however, voting

[27]This holds when the ratio of state-dependent payoffs $\frac{x_i}{y_i}$ is the same for all $i \in I$. Furthermore, this holds when $\pi_0 = \frac{1}{2}$ and $x_i + y_i$ is the same for all $i \in I$.

[28]A special case of this is a case in which z is known to the social planner.

rules less restrictive than simple majority are rarely observed. This section will show that in a setting with learning in which committee members can vote on how much they delay the decision, supermajority rule is indeed optimal.

Suppose that the planner has chosen a voting rule $\gamma < \frac{1}{2}$. Let us compare it to an alternative supermajority voting rule $\hat{\gamma} \equiv 1-\gamma > \frac{1}{2}$. As discussed in Section 2.3.3, when the decisions are different in different states, the committee will choose to learn the state. Let us look at different realisations of s. If $s < \gamma = 1-\hat{\gamma}$, then $g_{\gamma}(x) = g_{\hat{\gamma}}(x) = 0$ and $g_{\gamma}(y) = g_{\hat{\gamma}}(y) = 1$. Then under either voting rule the committee learns the state, and the reform is approved in state Y only. Hence, the ex ante expected welfare under either rule is $(1-\pi_0)\,\mathrm{E}(\mu_Y \mid s < \gamma)$. Similarly, if $s > 1-\gamma = \hat{\gamma}$, then $g_{\gamma}(x) = g_{\hat{\gamma}}(x) = 1$ and $g_{\gamma}(y) = g_{\hat{\gamma}}(y) = 0$. The committee again chooses to delay the decision until the state is revealed, and under either voting rule the reform is adopted in state X only. Then under either voting rule, the ex ante expected welfare is $\pi_0 \mathrm{E}(\mu_X \mid s > 1-\gamma)$.

What happens if s is between γ and $1-\gamma$? Under voting rule γ, this means that the reform is always adopted. In that case, the ex ante expected welfare equals $\mathrm{E}(z)$. On the other hand, under voting rule $\hat{\gamma}$, the reform is rejected in each state. Then, as discussed in Section 2.3.3, the committee will approve the reform at time 0 if $g_{\hat{\gamma}}(\pi_0 x + [1-\pi_0]\, y) = g_{\hat{\gamma}}(z) = 1$. This happens if and only if $z > 0$. Otherwise, the reform will never be approved. Thus, under $\hat{\gamma}$, the ex ante expected welfare when $s \in [\gamma, 1-\gamma]$ equals $\Pr(z > 0)\,\mathrm{E}(z \mid z > 0)$.

Hence, as long as $z > 0$, voting rule $\gamma < \frac{1}{2}$ and voting rule $\hat{\gamma} > \frac{1}{2}$ produce outcomes that are identical from the welfare point of view. If $z < 0$, then the outcomes are again welfare-equivalent unless s falls between γ and $1-\gamma$. In the latter case, the reform is rejected under $\hat{\gamma}$ and approved under γ. Therefore, the difference between expected utilitarian welfare of the committee members under the voting rule $\gamma < \frac{1}{2}$ and the utilitarian expected welfare under the voting rule $\hat{\gamma} = 1-\gamma > \frac{1}{2}$ is

$$\Pr(z < 0)\Pr(s \in [\gamma, 1-\gamma])\,\mathrm{E}(z \mid z > 0)$$

which is negative. Hence, expected welfare is larger under $\hat{\gamma}$ than under γ. This result is summarised the following proposition:

Proposition 8. *Suppose that $\lambda_X = \lambda_Y$ and that the committee cannot commit to rejecting the reform. Suppose further that all voters share the same ex ante payoff from the proposal, and that each voter gains from the proposal in one state and loses in the other. Let the share of voters who gain in state X be distributed with full support and independently of the ex ante expected payoff. Then for any voting rule $\gamma < \frac{1}{2}$, there exists*

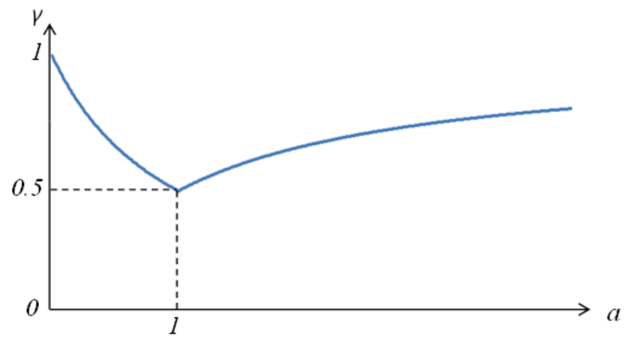

FIGURE 2.8: Optimal voting rule as a function of a.

an alternative rule $\hat{\gamma} > \frac{1}{2}$, such that the utilitarian welfare is higher under $\hat{\gamma}$ than under γ.

2.6.1 Specific Voting Rules

Consider the following restriction of the setting described above. Suppose that $\pi_0 = \frac{1}{2}$, and that for each voter, $(x_i, y_i) \in \{(a, -1), (-1, a)\}$. Thus, as before, each voter either gains a or loses 1 from the reform. The parameter a measures the magnitude of gains relative to that of losses. Note that the expected payoff from the reform equals $z = \frac{1}{2}(a-1)$, for all i.

Suppose that a (and hence z) is known to the social planner, but s, the share of agents with preferences $(a, -1)$, is unknown. Let s be drawn from some distribution F with full support and with the associated density f. Which voting rule γ should the designer select to maximise utilitarian welfare? Denote the welfare-maximising voting rule by γ^*. The following proposition provides the expression for it:

Proposition 9. *If $\pi_0 = \frac{1}{2}$ and $(x_i, y_i) \in \{(a, -1), (-1, a)\}$ $\forall i \in I$, then for any F, the optimal voting rule is given by:*

$$\begin{cases} \gamma^* = \frac{1}{1+a} & \text{if } a < 1 \\ \gamma^* = \frac{a}{1+a} & \text{if } a > 1 \end{cases}$$

Proof. See Appendix.

Figure 2.8 shows the optimal voting threshold as a function of a. As we can see, the optimal voting rule is a simple majority rule when $a = 1$, i.e. when individual gains and losses from the reform have the same magnitude. Otherwise, it increases as a moves away from 1 in either direction. Thus, greater supermajority is optimal when individual gains are much larger or much smaller than losses.

The intuition behind this result is the following. If a is close to zero, then it is optimal to reject the reform even if only a few voters lose from it, because their losses are much larger than the gains of winners. Hence, a large supermajority is optimal. If a is very large, then it is only optimal to reject the reform if there number of winners is very small. But $a > 1$ means that the ex ante expected payoff from the reform is positive for all voters. Then, making γ large means that the reform is likely to be rejected in either state - and if $a > 1$, this means that all voters agree to adopt the reform in ignorance at time 0. Thus, making the voting threshold larger *increases* the likelihood that the reform is adopted, which is exactly what a benevolent social planner wants to achieve when $a > 1$.

Now suppose that the social planner not only does not know s, but also has no knowledge of a, and hence does not know the expected payoff from the reform. This reflects settings in which the planner needs to select an optimal constitutional rule for a committee that has to vote on a number of reforms that may differ in their expected social value.

Suppose that a is drawn from some distribution H with support on $[0, +\infty)$. For simplicity, I assume that this distribution has no mass points at $a = 1$ or at $a = 0$. As before, s is drawn from some distribution F, and is independent of a. Each agent is informed about the realisations of a and s before voting begins, but the planner does not know them when choosing γ.

Proposition 10. *If $\pi_0 = \frac{1}{2}$ and $(x_i, y_i) \in \{(a, -1), (-1, a)\}\ \forall i \in I$, then for any F and H, the optimal voting rule is given by:*

$$\gamma^* = \min\left\{\frac{1 + \Pr(a > 1)\left[\mathrm{E}(a \mid a > 1) - 1\right]}{\mathrm{E}(a) + 1}, 1\right\}$$

Proof. See Appendix.

How does the optimal decision rule depend on the distribution of a? If a is very likely to be large, then $\Pr(a > 1) \approx 1$, and $\mathrm{E}(a \mid a > 1) \approx \mathrm{E}(a)$. Thus, $\gamma^* \approx \frac{\mathrm{E}(a)}{\mathrm{E}(a)+1}$. This is a more general version of the result in Proposition 9 for the case when $a > 1$. On the other hand, if $\Pr(a > 1) = 0$, then $\gamma^* = \frac{1}{\mathrm{E}(a)+1}$, which is the a version of the result in Proposition 9 for the case when $a < 1$.

Now suppose that a is likely but not certain to be below 1, so $\Pr(a > 1)$ is positive but very small. If $\mathrm{E}(a \mid a > 1)$ is also relatively small, then $\gamma^* \approx \frac{1}{\mathrm{E}(a)+1}$, which is similar to the case when $\Pr(a > 1) = 0$. Suppose, however, that the expected gains from the reform are low relative to losses, but there is a tiny probability that the gains can be very large. In terms of the above expression, this means that $\Pr(a > 1)$ is positive but very small, while $\mathrm{E}(a \mid a > 1)$ is large. Then, for a sufficiently large $\mathrm{E}(a \mid a > 1)$, the optimal voting rule becomes close to unanimity (or equals unanimity).

Hence, if gains from the reform are very likely to be lower than losses, then a small probability of very large gains has a disproportionate effect on the optimal decision rule. To put it differently, the optimal decision rula is disproportionately affected by the shape of the right tail of the distribution of gains.

Finally, suppose that $\mathrm{E}(a) = 1$ and that the median of a is 1.[29] What happens to the optimal voting rule if uncertainty about payoffs increases? Suppose that we impose a mean- and median-preserving spread of the distribution of a. Then $\mathrm{E}(a)$ and $\Pr(a > 1)$ remain the same, but $\mathrm{E}(a \mid a > 1)$. It is clear from the expression above that γ^* becomes larger.

In other words, greater ex ante uncertainty about payoffs makes the optimal decision rule more conservative. This result is somewhat counterintuitive, since neither voters as well nor the social planner are risk-neutral. Hence, if expecter gains and losses from the reform remain the same, greater uncertainty should not by itself call for a more conservative decision rule. The reason why it does is that, when information is not acquired, the proposal is approved if and only if $a > 1$. Imposing a spread on the distribution of a increases the expected value of a conditional on a being above the median. Hence, the expected welfare conditional on the proposal being adopted in ignorance becomes larger. This increases the social value of ignorance, and a benevolent social planner responds by making it more likely that the committee opts for ignorance. Hence, she sets the optimal threshold higher.

2.7 Commitment to Learning

In addition to the choice of a voting rule, there is another way in which institutional settings can affect the collective decision. Namely, the decision-making body can commit to learning the state regardless of the collective decision. For example, the legislative process often requires parliaments to have several readings before a law is passed, and overriding this requirement is often hard or impossible. A strong tradition of public debate or a constitutional guarantee of transparency or freedom of speech can also serve

[29] For example, suppose that $a \in [0, 2]$, and is distributed symmetrically on that interval.

as commitment devices that ensure learning the information relevant to a collective decision. When is such a commitment optimal?

Consider the case when $\lambda_X = \lambda_Y$, and $\gamma = \frac{1}{2}$. Suppose that the committee has a power to commit to rejecting a proposal without learning. A social planner is considering to impose a requirement that the committee must learn the state before adopting the reform. Suppose that she judges outcomes based on a welfare function $w : \mathbb{R}^I \to \mathbb{R}$ which maps expected payoffs of individuals (given the information available to the planner) to social welfare. As a normalisation, suppose $w(0, 0, ...) = 0$. Let $sign(a)$ be the sign (positive or negative) of a scalar a. To simplify notation, denote $d(z) \equiv g_{0.5}(z) - \frac{1}{2}$, so that given a vector of payoffs $z \in \mathbb{R}^I$, the reform is adopted whenever when $d(z)$ is positive.

Proposition 11. *Commitment to learning is weakly socially preferable if* $sign\left[d\left(\pi_0 x + [1-\pi_0] y\right)\right] \neq sign\left[w\left(\pi_0 x + [1-\pi_0] y\right)\right]$*, and is weakly harmful if* $sign\left[d\left(\pi_0 x + [1-\pi_0] y\right)\right] = sign\left[w\left(\pi_0 x + [1-\pi_0] y\right)\right]$

Proof. See Appendix.

Intuitively, this proposition says that information release is weakly preferable whenever the decision that the committee makes in ignorance is different from the welfare-maximising decision.

Suppose that the welfare function is utilitarian, i.e. equals the sum of payoffs. Then information release is optimal when the distribution of ex ante expected payoffs $\pi_0 x_i + [1-\pi_0] y_i$ across players has a mean and a median that are of different signs. Referring to Figure 2.5 above, this can happen when the distribution of payoffs is skewed along the "Southwest-Northeast" axis. This happens, for example, when the majority of agents benefit from the proposal in expectation (making the median positive), but there is a minority of individuals who each lose much if the proposal is accepted (resulting in a negative mean). Hence, a constitutional guarantee of transparency can serve as a mechanism to protect a minority, and it is optimal when there is minority with a large stake in the outcome of the vote.

2.8 Extensions and Robustness

2.8.1 Imperfect Signals

The setting in which $\lambda_X = \lambda_Y$ is equivalent to a one-shot model in which the group faces a choice between learning the state prior to voting on the reform, or voting in ignorance. We can, however, consider a more general static model model in which the signal that the group can acquire is imperfect. This section will show that the results obtained in the previous section easily generalise in this setting.

Suppose that, similar to the baseline model above, the group can choose to wait and costlessly acquire a binary signal $\sigma \in \{X, Y\}$. Let $\Pr(\sigma = X \mid \omega = X) = p$ and $\Pr(\sigma = X \mid \omega = Y) = q$, where $p > q$. Thus, if signal X arrives, the posterior probability that the state is X increases relative to π_0; and if signal Y arrives, it decreases relative to π_0.

If the committee chooses to acquire information, then, upon receiving signal X, they will believe that the state is X with probability $\frac{\pi_0 p}{\pi_0 p + (1-\pi_0) q}$. In this case, voter i's expected payoff if the project is approved is $\frac{\pi_0 p}{\pi_0 p+(1-\pi_0)q} x_i + \frac{\pi_0(1-\pi_0)q}{\pi_0 p+(1-\pi_0)q} y_i$. Thus, when signal X is received, the project will be approved if and only if

$$g\left[\frac{\pi_0 p}{\pi_0 p + (1-\pi_0) q} x_i + \frac{\pi_0 (1-\pi_0) q}{\pi_0 p + (1-\pi_0) q} y_i\right] = 1$$

or equivalently, if and only if $g_\gamma [\pi_0 p x + (1-\pi_0) q y] = 1$. Similarly, if signal Y is received, the posterior probability that the state is X equals $\frac{\pi_0(1-p)}{\pi_0(1-p)+(1-\pi_0)(1-q)}$. Then voter i gains $\frac{\pi_0(1-p)}{\pi_0(1-p)+(1-\pi_0)(1-q)} x_i + \frac{(1-\pi_0)(1-q)}{\pi_0(1-p)+(1-\pi_0)(1-q)} y_i$ in expectation if the proposal goes ahead - hence, the proposal goes ahead if and only if

$$g_\gamma\left[\frac{\pi_0 (1-p)}{\pi_0 (1-p) + (1-\pi_0)(1-q)} x + \frac{(1-\pi_0)(1-q)}{\pi_0 (1-p) + (1-\pi_0)(1-q)} y\right] = 1$$

or equivalently, if and only if $g_\gamma [\pi_0 (1-p) x + (1-\pi_0)(1-q) y] = 1$.

Ex ante, if information is not acquired, agent i's expected payoff from adopting the proposal is $\pi_0 x_i + (1-\pi_0) y_i$. Then in ignorance, the group adopts the proposal whenever $g_\gamma [\pi_0 x_i + (1-\pi_0) y_i] = 1$.

Suppose that the group has an option to commit to reject the proposal. Proposition 12 below shows that the necessary and sufficient condition for the committee to have a collective preference for ignorance is similar to the case with perfect signals described in Section 3. Specifically, the group will weakly prefer not to acquire information if and only if the collective decision upon observing signal X is different from the collective decision upon observing signal Y. Moreover, the group will strictly prefer not to acquire information if and only if the collective decisions upon receiving the two signals are the same, and both are different from the collective decision made at the initial belief π_0.

Proposition 12. *The group has a weak preference for ignorance if and only if* $g[\pi_0 p x + (1-\pi_0) q y] = g[\pi_0 (1-p) x + (1-\pi_0)(1-q) y]$. *Furthermore, the committee has a strict preference for ignorance if and only if* $g[\pi_0 p x + (1-\pi_0) q y] = g[\pi_0 (1-p) x + (1-\pi_0)(1-q) y] \neq g[\pi_0 x_i + (1-\pi_0) y_i]$.

Proof. See Appendix.

2.8.2 Generic Information Structure

The one-shot learning model described above has so far assumed that the state space is binary. A more general approach would be to see the state of the world as an element of a generic set. Consider a mapping from this set of states to some set of messages - in other words, an information partition. The learning decision can then be described as a decision to replace the prior belief with this partition.

Suppose that there is a finite set of states Ω. Each state $j \in \Omega$ occurs with a prior probability p^j, which is common knowledge. If the reform is approved, then in state j each agent $i \in I$ receives payoff x_i^j - throughout this section, subscripts will denote agents, while superscripts will denote states. Let x^j be a vector representing the payoffs of all agents if the proposal is approved and the state is j. The committee first chooses whether to acquire information partition $\mathcal{P}$, which is a partition of Ω. Let us denote by S a generic element of $\mathcal{P}$; I will refer to S as a *message.*

Suppose that decisions are made by simple majority. If the commitee decides against learning, then the decision on the reform is based on the prior. Hence, it will approve the project if and only if $g_{0.5}\left[\sum\limits_{j\in\Omega} p^j x^j\right] = 1$. Then ex ante, if the committee chooses to stay ignorant, agents will receive the expected payoff vector $\sum\limits_{j\in\Omega} p^j x^j g_{0.5}\left[\sum\limits_{j\in\Omega} p^j x^j\right]$.

Now suppose that the partition $\mathcal{P}$ is acquired. If the committee receives a message $S \in \mathcal{P}$, the posterior probability that the state is j will, by Bayes law, be $\frac{p^j}{\Pr(S)}$ if $j \in S$, and zero otherwise, where $\Pr(S)$ denotes the prior probability of receiving the message S. Then, upon receiving the message S, the committee will vote in favour of the proposal iff $g_{0.5}\left[\sum\limits_{j\in S} \frac{p^j x^j}{\Pr(S)}\right] = 1$. The ex ante expected payoff vector to all agents will then equal

$$\sum_{S\in\mathcal{P}}\left(\Pr(S)\, g_{0.5}\left[\sum_{j\in S}\frac{p^j x^j}{\Pr(S)}\right]\sum_{j\in S}\frac{p^j x^j}{\Pr(S)}\right) = \sum_{S\in\mathcal{P}}\left(g_{0.5}\left[\sum_{j\in S} p^j x^j\right]\sum_{j\in S} p^j x^j\right)$$

where the last equality uses the fact that $g_\gamma(\cdot)$ is invariant under a rescaling of payoffs.

To avoid the awkward case in which every agent is indifferent between acquiring and not acquiring information , I will assume that information is *non-trivial.* Specifically, I will assume that there exists an $S \in \mathcal{P}$ such that $g_{0.5}\left[\sum\limits_{j\in S}\frac{p^j x^j}{\Pr(S)}\right] \neq g_{0.5}\left[\sum\limits_{j\in\Omega} p^j x^j\right]$. In words, there is at least one message that induces a decision different from the one that is made without information. Hence, information can potentially have some effect.

Then the following result can be derived:

Proposition 13. *Assume that information is non-trival and* $\gamma = \frac{1}{2}$. *Then the information partition* $\mathcal{P}$ *will be acquired iff* $g_{0.5}\left[\sum_{j \in S \in M} p^j x^j\right] \neq g_{0.5}\left[\sum_{j \in \Omega} p^j x^j\right]$, *where* M *is the collection of all* $S \in P$ *for which* $g_{0.5}\left[\sum_{j \in S} p^j x^j\right] \neq g_{0.5}\left[\sum_{j \in \Omega} p^j x^j\right]$.

Proof. See Appendix.

Proposition 13 says the following. Take all the messages in the information structure $\mathcal{P}$ that induce a decision different from the one made without information. Now suppose that the committee only knows that one of such messages will be received, without knowing which one. If, given this knowledge, the committee still makes the same decision (i.e. a decision different from the one they make in ignorance), then the committee will vote to acquire information structure $\mathcal{P}$.

Note that the assumption of information being non-trivial implies that M is non-empty.

We can compare this result to the case of the binary state. If M includes only one state (which happens when $g_{0.5}(x) \neq g_{0.5}(y)$), the condition in Proposition 13 is satisfied automatically, and the committee chooses to learn the state. If M includes both states (this happens when $g_{0.5}(x) = g_{0.5}(y) \neq g_{0.5}(\pi_0 x + [1 - \pi_0] y)$), then the condition in Proposition 13 automatically fails, and information about the state is not acquired.

2.8.3 Limited Decision Window

The preceding analysis assumed that the vote on the reform can be made at any time. Suppose instead that the committee can only vote on the reform at a closed subset $S \subset [0, +\infty)$ of time. Two distinct cases are worth examining: the case when S is bounded from above (for example, when the reform can only be adopted before a certain deadline), and the case when it is not bounded (for example, when the committee is only able to meet at certain intervals of time).

When S is not bounded, then the analysis of the case when $\lambda_X = \lambda_Y$ remains essentially the same. In a Markov-perfect equilibrium, the committee can either vote on the reform at $t = 0$, or at the earliest time that belongs to S but is after the state is revealed. Thus, just as in the preceding analysis, the committee's choice is between voting on the reform at the prior belief, or waiting until the state is revealed. Hence, the analysis proceeds as before, and the conclusions of Propositions , 1, 2, 3, 8, 11, and 12 still hold.

Suppose S is bounded from above, and let $\hat{t} \equiv \max(S)$. Consider the case when the committee has the power to commit to rejecting the reform. If the committee is waiting then, as before, the belief π_0 does not change as long as the state is not revealed. Supposed the state is not revealed until $\hat{t}$. Then at $\hat{t}$, the committee's choice is between stopping and voting on the reform, and foregoing the reform. The committee will stop and adopt the reform at $\hat{t}$ whenever $g_\gamma(\pi_0 x + [1-\pi_0] y) = 1$. Otherwise, the reform is rejected (either because the committee votes and rejects it, or because it continues waiting past $\hat{t}$). Given $\hat{t}$ and the rate of learning, let $\mu \in (0,1)$ denote the probability that the state is revealed at or before $\hat{t}$. Then at $t = 0$, the payoff of agent i if the committee decides to wait equals

$$\mu(\pi_0 x_i g_\gamma(x) + [1-\pi_0] y_i g_\gamma(y)) + (1-\mu)(\pi_0 x_i g_\gamma(x) + [1-\pi_0] y_i g_\gamma(y)) g_\gamma(\pi_0 x + [1-\pi_0] y)$$

while if the committee decides to vote on the proposal at $t = 0$, i's payoff is $(\pi_0 x_i + [1-\pi_0] y_i) g_\gamma(\pi_0 x + [$ Then i's gain from choosing not to learn and to vote on the reform at $t = 0$ equals

$$\mu(\pi_0 x_i + [1-\pi_0] y_i) g_\gamma(\pi_0 x + [1-\pi_0] y) - \mu(x_i g_\gamma(x) + y_i g_\gamma(y)) = \mu d_i$$

where d_i is the gain from ignorance for agent i as defined in Section 2.3.1. The committee opts for ignorance whenever $g_\gamma(\mu d) = 1$, where Since $g_\gamma(\cdot)$ is invariant to rescaling of payoffs, $g_\gamma(\mu d) = 1$ is equvialent to $g_\gamma(d) = 1$, which is the same condition as in Section 2.3.1. Hence, the necessary and sufficient condition for the committee to choose not to learn is the same as before, and all the analysis proceeds in the same way.

We can thus conclude that the key results of the chapter are robust to restricting the time window of the collective decision on reform.

2.8.4 Choice Between Learning Processes

Consider again the case when the two states are revealed at the same rate. Recall that this setting is equivalent to a model in which the committee makes a one-shot decision between voting on the reform in ignorance and waiting until the state is revealed. Suppose that instead of that choice, the committee is facing a choice between a more informative and a less informative learning process. That is, there is a process 0 which reveals the true state at a rate λ_0, and a process 1 which reveals it at a rate $\lambda_1 > \lambda_0$. The committee can vote to use the former, less informative process instead of the latter one if the share of agents who are in favour of this is at least γ.

Choosing a less informative learning process is similar to choosing to make the decision in ignorance. The following shows a committee that votes against ignorance will also vote against selecting a less informative process.

If the committee can make its decision at any point in time, the choice of a learning process is irrelevant - under each process, the committee will, depending on the distribution of preferences, either wait until the state is revealed or vote in ignorance. Same happens if the decision can only be made at times $t \in S$ for some unbounded S.

Suppose, however, that S is bounded (and closed). For a process $j \in \{0, 1\}$, let

$$\mu_j \equiv 1 - \mathrm{e}^{-\lambda_j \hat{t}}$$

be the probability that the state is revealed before the deadline $\hat{t}$. Then under process 0, the committee votes on the reform at $t = 0$ whenever $g_\gamma(\mu_0 d) = 1$, and under process 1 it makes its decision at $t = 0$ whenever $g_\gamma(\mu_1 d) = 1$. As $g_\gamma(\cdot)$ is invariant to rescaling of payoffs, these two conditions are equivalent to $g_\gamma(d) = 1$. Thus, if $g_\gamma(d) = 1$, under either process the committee makes the decision in ignorance, and individuals are indifferent as to which process is chosen.

If $g_\gamma(d) = 0$, then the committee chooses to wait until the the state is revealed. If the state is not revealed by the deadline $\hat{t}$, then at $\hat{t}$ the belief about the state is still π_0. The committee then adopts the reform at $\hat{t}$ whenever $g_\gamma(\pi_0 x + [1 - \pi_0] y) = 1$. Hence, given a learning process $j \in \{0, 1\}$, $i's$ expected payoff when $g_\gamma(d) = 0$ is

$$\mu_j (\pi_0 x_i g_\gamma(x) + [1 - \pi_0] y_i g_\gamma(y)) + (1 - \mu_j)(\pi_0 x_i + [1 - \pi_0] y_i) g_\gamma(\pi_0 x + [1 - \pi_0] y)$$

Then the gain in i's expected payoffs if the committee switches from learning process 1 to a less informative learning process 0 is

$$(\mu_1 - \mu_0) [(\pi_0 x_i + [1 - \pi_0] y_i) g_\gamma(\pi_0 x + [1 - \pi_0] y) - \pi_0 x_i g_\gamma(x) - [1 - \pi_0] y_i g_\gamma(y)]$$

which equals

$$(\mu_1 - \mu_0) d$$

Since $g_\gamma(d) = 0$, the committee never switches to process 0. Thus, as long as the committee prefers learning the state, it will also prefer a more informative learning process.

2.9 Conclusions

The aim of this chapter was to analyse a committee's choice between learning the state of the world and not learning it, prior to voting on a reform. Information was assumeed to be costless and public. Because information can change the eventual collective decision, some committee members may be against acquiring information, and under some conditions, the share of these members may be enough for the committee to choose ignorance.

It turned out that this choice depends on the group's decision in different states. When the decision on the reform is the same in two states, the committee weakly prefers to stay uninformed. When these decisions are, additionally, different from the decision made without information, the preference for ignorance is strict.

Specific types of payoff distributions induce a collective preference for ignorance. The committee will choose to remain uninformed if and only if the group is more fractionalised on the state-relevant dimension than on the state-irrelevant dimension. Thus, groups that are sufficiently heterogeneous with respect to their attitude towards the reform will choose to vote on the reform without learning the state.

When two states are revealed at a different rates, voters whose preferred state is revealed slower than their least preferred state are better off. Hence, when the decision is made by a group consisting of agents with opposing preferences, there is a benefit of learning slower.

A further result is that changing the plurality voting rule has a non-monotone effect on collective decisions. If the minimum number of votes required to stop waiting and adopt the reform is reduced, this at first expands the range of beliefs at which the committee decides to stop learning and accept the reform. At some point, however, the decision threshold becomes so low that the committee switches to gathering the maximum amount of information.

It was also found that under certain restrictions on the distribution of preferences, any voting threshold below 0.5 is dominated from the welfare point of view by a supermajority rule. This may partly explain why supermajority rules are often used, while plurality rules requiring less than simple majority are not.

This work also looked at welfare effects of making collective learning compulsory. Such a committment to transparency is optimal when the decision made in ignorance is different from a welfare-maximising decision - which, in the case of majority voting and utilitarian welfare criterion, happens when the distribution of average payoffs across states is skewed.

This implies that a committment to transparency can be an instrument for protecting the interests of a minority with a large stake in the decision.

The result if members of a group are fractionalised on their attitude to consequences of a decision, the group is less likely to choose to learn these consequences implies several testable predictions. For example, we can expect to see less public discussion and less demand for expert advice in societies in which there is less consensus over desirable effects of policies. Support for public debate and freedom of information can be lower in countries with more fractionalised electorates. Decisions on divisive issues are likely to be made in haste. Subsequent research can thus focus on testing these predictions empirically or experimentally.

Chapter 3

Sincere Voting as a Strategic Equilibrium

3.1 Introduction

In 1785, Marquis de Condorcet looked at a group of voters attempting to reach a correct decision. Each voter has imperfect information about which decision is correct. Condorcet's famous result, known as the Condorcet Jury Theorem, states that as the size of the group increases, so does the probability of them reaching the correct decision[1].

This celebrated result relies on an assumption that every voter votes "sincerely", i.e. follows his private information and acts as if he was pivotal. More recently, this assumption has been challenged. Austen-Smith and Banks (1996) and Feddersen and Pesendorfer (1997) have looked at equilibria in voting games when each voter receives an imperfect signal about a payoff-relevant state of the world. In these situations, each agent votes to maximise her expected payoff conditional on her signal and on the event that her vote is pivotal. In such a setting, sincere voting is not, in general, an equilibrium strategy. In fact, in large elections, only a vanishingly small proportion of voters follow their private signal, even though voting aggregates information. This result is surprising, since there is not much evidence that voters in real-life elections do vote against the information they have[2].

This paper aims to provide game-theoretic foundations for the intuitive hypothesis that voters vote sincerely. It models a situation in which voters need to choose between adopting or rejecting a certain proposal, called "reform". The payoff of every voter from

[1] See Condorcet (1976).

[2] Degan and Merlo (2009) provide some evidence than they do not.

that reform depends on a binary state of the world, and each voter receives an imperfect binary signal about the state.

The key difference between this model and much of the literature is that voters' I do not assume that voters share similar preferences[3]. Specifically, all voters are divided into two groups. Members of the first group gain from the reform in the first state but lose in the second; while members of the second group gain in the second state but lose in the first. We can refer to these two groups as a majority and a minority.

There are a number of real-world situations in which such a distribution of preferences can be observed. For example, suppose that a referendum is held in some country on an economic reform, such as trade liberalisation. It is believed that the reform can either benefit manufacturing firms but hurt agriculture, or it may help agriculture while damaging the manufacturing sector[4]. The effect of liberalisation can be called a state of the world. Voters who are employed in agriculture prefer the reform to be adored if and only if it benefits agriculture, while voters who are affiliated with manufacturing are in favour of the reform if and only if manufacturing gains.

As a further example, consider an election in which voters need to decide whether to re-elect an incumbent or to elect a challenger. The incumbent's position is well-known, but there is some uncertainty about the policies that the challenger will implement if she is elected. It might be that the challenger's policies will be more left-wing than those of the incumbent, or it might be that challenger is more right-wing than the incumbent. Left-wing voters prefer the challenger to be elected in the former but not in the latter case, while right-wing voters have the opposite preferences.

The key result of the paper, established in Proposition 16, is an equilibrium in which every voter acts sincerely exists under certain ranges of parameters. In particular, a sincere voting equilibrium exists when the minority is close in size to the majority. It also exists when the voting rule is sufficiently close to simple majority, well as when signals are sufficiently precise. These results hold for an arbitrary size of the voting group, and not just for large electorates, which have been at the centre of attention of much of the literature.

Two previous papers - namely, Damiano et al. (2011) and Acharya and Meirowitz (2014) - have also provided conditions under which sincere voting is an equilibrium strategy.

[3]Much of the prior literature looks at common value settings, in which changing the state from one to another changes the utility of all voters from a given alternative in the same direction. A frequently cited example is voting in a jury, in which all members want to acquit the defendant if he is innocent, and convict him if he is guilty (though the intensity of preferences can differ).

[4]The fact that trade liberalisation and other economic reforms can create winners and losers, and that identities of winners and of losers are often ex ante uncertain has received much attention in literature on development and growth. See, for example Fernandez and Rodrik (1991), Rodrik (1993).

In Damiano et al. (2011), sincere voting is driven by the possibility of a recount if the voting outcome is sufficiently close. In Acharya and Meirowitz (2014), sincere voting by informed voters can be an equilibrium if uninformed voters are present. Both papers deal with large elections, in which the size of the electorate approaches infinity.

This paper is also related to Bhattacharya (2013*b*), who also considers information aggregation in large elections in which preferences are non-monotone in the state[5]. However, Bhattacharya (2013*b*) does not focus on the conditions under which agents vote sincerely.

3.2 Model

Consider a group of $n+1$ voters who need to decide whether to approve or reject a reform. Each voter i has a type $x_i \in \{0,1\}$, which represents her preferences. There is an unknown state of the world $y \in \{0,1\}$. If the reform is rejected, each voter receives a payoff of zero. If the reform is approved, voter i receives a payoff $u(x_i, y)$. Let $u(x_i, y) = a > 0$ if $x_i = y$; and $u(x_i, y) = -1$ if $x_i \neq y$. Thus, each voter wins from the reform if the state corresponds to her type, and loses otherwise; and and a measures the magnitude of gains relative to that of losses.

In the beginning of the game, Nature selects the state; each state is selected with an equal probability. Then Nature draws the voter types; each type is drawn independently, and the probability that a voter's type is 1 is $\gamma \in (0,1)$. Then, each voter receives a private signal $s_i \in \{0,1\}$ about the state. The realisation of the signal corresponds to the true state with probability $p \geq \frac{1}{2}$. Thus, p measures the accuracy of the signal. After observing her signal, each voter decides whether to vote for or against the reform. The reform is adopted if the number of votes in favour of it is strictly larger than some number k. The size of k thus represents the type of the majority rule that is used. For instance, $k = \frac{n}{2}$ represents a simple majority rule, while $k = n$ represents unanimity rule. After the vote is made, payoffs are realised depending on whether the reform is adopted and on the state.

Denote by $g_i(x_i, s_i) \in [0,1]$ each voter's action (the probability of voting for the reform) at the equilibrium as a function of her signal and her type. Let $\hat{g}_i(x_i, s_i)$ be each voter's sincere action.

[5] See also Kim and Fey (2007).

3.3 Analysis

The subsequent analysis will first examine voting outcomes when all agents vote sincerely. Then, I will derive the equilibrium when agents vote strategically, and find when the two correspond.

3.3.1 Sincere Voting

Suppose that every agent votes as if she were pivotal. Then each voter backs the reform if and only if

$$\mathrm{E}\left[u\left(x_i, y\right) \mid s_i\right] > 0$$

Thus, a voter for whom $s_i = x_i$ votes in favour of the reform if and only if $pa - (1-p) > 0$. A voter for whom $s_i \neq x_i$ votes in favour of the reform if and only if $-p + (1-p)\,a > 0$. Thus, we have the following result:

Proposition 14. *If each voter votes sincerely, then all voters follow a symmetric strategy* $\hat{g}\left(x_i, s_i\right)$*, of the following form:*

- *When* $x_i = s_i$, $\hat{g}\left(x_i, s_i\right) = 1$ *iff* $\frac{1-p}{p}\frac{1}{a} \leq 1$
- *When* $x_i \neq s_i$, $\hat{g}\left(x_i, s_i\right) = 1$ *iff* $\frac{p}{1-p}\frac{1}{a} \leq 1$

If $\frac{1-p}{p}\frac{1}{a} = 1$ or if $\frac{p}{1-p}\frac{1}{a} = 1$, some voters may be indifferent between voting for or against the reform. To avoid cumbersome notation, I will abstract from these special cases and assume that no voter is indifferent.

3.3.2 Strategic Voting

Suppose that each voter behaves strategically. Then she knows that her vote can only affect the outcome if she is pivotal - i.e. if exactly k other voters have voted in favour of the reform. Denote this event by *piv*. If this event happens with a positive probability, a voter i who is strategic will vote to support the reform if and only if

$$\mathrm{E}\left[u\left(x_i, y\right) \mid s_i\,,\, piv\right] \geq 0$$

I will focus on symmetric strategies $g\left(x_i, s_i\right)$. Let q_0 and q_1 denote the probabilities (unconditional on signals) that a voter votes in favour of the reform when the true state is respectively, 0 and 1. These probabilities, of course, depend on voters' equilibrium strategies. When $y = 0$, each voter is pivotal with probability $\left(\begin{array}{c} k \\ n \end{array}\right) q_0^k (1-q_0)^{n-k}$; and when $y = 1$, a voter is pivotal with probability $\left(\begin{array}{c} k \\ n \end{array}\right) q_1^k (1-q_1)^{n-k}$.

Before proceeding with the bulk of the analysis, let us check what happens if q_0 equals either zero or one. Note that a voter receives each signal $s_i \in \{0,1\}$ with probability $\frac{1}{2}$ and can have each type $x_i \in \{0,1\}$ with a strictly positive probability. Therefore, $q_0 = 0$ or $q_0 = 1$ can only happen if $g(x_i, s_i)$ is the same regardless of x_i or s_i. In that case, $q_1 = q_0$.[6]

There are two cases in which each voter is pivotal with probability one. First, it is possible that $q_1 = q_0 = 0$ and $k = 0$. Proposition 14 implies that this happens when $a \geq \frac{p}{1-p} \geq \frac{1-p}{p}$. Second, it is possible that $q_1 = q_0 = 1$ and $k = n$. From Proposition 14, this requires that $a \leq \frac{1-p}{p} \leq \frac{p}{1-p}$. In these cases, sincere voting is a strategic equilibrium. Intuitively, this holds when when p is close to 0.5 (i.e. signals are uninformative), and a is far from 1 (so ex ante, each voter gains or loses much from the reform). In these cases, a voter has no strategic incentive to condition her votes on her signal.

When q_0 and q_1 are both zero or one, and k is distinct from n and from 0, then no voter is ever pivotal. In that case, no voter has an incentive to deviate, so this is an equilibrium.

Consider the case in which $q_1, q_0 \in (0,1)$. Then each voter is pivotal with a positive probability. The strategies of all voters are then described by the following proposition, the proof of which is given in the Appendix:

Proposition 15. *At the equilibrium in which each voter is pivotal with a positive probability, all voters play a strategy $g(x_i, s_i)$, of the following form:*

- $g(0,0) = 1$ *iff* $\frac{p}{1-p} a \geq M$
- $g(1,0) = 1$ *iff* $\frac{p}{1-p} \frac{1}{a} \leq M$
- $g(0,1) = 1$ *iff* $\frac{1-p}{p} a \geq M$
- $g(1,1) = 1$ *iff* $\frac{1-p}{p} \frac{1}{a} \leq M$

where

$$M \equiv \frac{q_1^k (1-q_1)^{n-k}}{q_0^k (1-q_0)^{n-k}}$$

[6]Through the same logic, $q_1 = 0$ or $q_1 = 1$ implies that $q_0 = q_1$.

3.3.3 Sincere Voting as a Strategic Equilibrium

In this section I will derive some sufficient conditions under which at the equilibrium every voter acts as if she were voting sincerely. That happens when conditions in Proposition 14 match those in Proposition 15.

If $a \geq \frac{p}{1-p} \geq \frac{1-p}{p}$, or if $a \leq \frac{1-p}{p} \leq \frac{p}{1-p}$, then sincere voting is an equilibrium, as explained above. Let us now focus on the more interesting case, in which $a \in \left(\frac{1-p}{p}, \frac{p}{1-p}\right)$.

In that case, if voting is sincere, then $\hat{g}(0,0) = \hat{g}(1,1) = 1$, and $\hat{g}(0,1) = \hat{g}(1,0) = 0$. Intuitively, when gains and losses are relatively close in magnitude and signals are relatively precise, then sincere voting implies supporting the reform when the voters' signal corresponds to her type.

We need to establish conditions under which voters' equilibrium actions mimic their sincere action - i.e. the conditions under which $g(\cdot,\cdot) = \hat{g}(\cdot,\cdot)$. From Proposition 15, $g(0,0) = g(1,1) = 1$ implies that $\frac{1-p}{p}\frac{1}{a} \leq M \leq \frac{p}{1-p}a$. Similarly, $g(0,1) = g(1,0) = 0$ implies that $\frac{1-p}{p}a \leq M \leq \frac{p}{1-p}\frac{1}{a}$.[7] Putting the two conditions together and using the expression for M, we can see that a strategic equilibrium in which all voters vote sincerely exists when:

$$\max\left\{\frac{1-p}{p}\frac{1}{a}, \frac{1-p}{p}a\right\} \leq M \leq \min\left\{\frac{p}{1-p}\frac{1}{a}, \frac{p}{1-p}a\right\}$$

Recall that $M = \left(\frac{q_1}{q_0}\right)^k \left(\frac{1-q_1}{1-q_0}\right)^{n-k}$. However, q_0 and q_1 are also equilibrium objects. At the equilibrium in which all voters act sincerely, when the true state is 0, a voter supports the reform if her type is 0 and she receives a correct signal; or if her type is 1 and she observes the wrong signal. Thus, $q_0 = (1-\gamma)p + \gamma(1-p)$. Similarly, when $y = 1$, a voter votes in favour of the reform if her type is 1 and she gets the right signal; or if her type is 0 and she receives the wrong signal. Thus, $q_1 = \gamma p + (1-\gamma)(1-p) = 1 - q_0$. Hence, if all voters act sincerely, we have

$$M = \left(\frac{1-q_0}{q_0}\right)^{2k-n} = \left[\frac{\gamma p + (1-\gamma)(1-p)}{(1-\gamma)p + \gamma(1-p)}\right]^{2k-n}$$

and the sufficient condition for the existence of a sincere voting equilibrium is

$$\max\left\{\frac{1-p}{p}\frac{1}{a}, \frac{1-p}{p}a\right\} \leq \left[\frac{\gamma p + (1-\gamma)(1-p)}{(1-\gamma)p + \gamma(1-p)}\right]^{2k-n} \leq \min\left\{\frac{p}{1-p}\frac{1}{a}, \frac{p}{1-p}a\right\} \tag{3.1}$$

[7] Note that $a \in \left(\frac{1-p}{p}, \frac{p}{1-p}\right)$ implies $\frac{1-p}{p}\frac{1}{a} < \frac{p}{1-p}a$ and $\frac{1-p}{p}a < \frac{p}{1-p}\frac{1}{a}$.

Note that when the size of the group becomes arbitrarily large - i.e. when $n \to \infty$ and k remains a constant fraction of n - then M converges to zero or infinity (except for the special case in which $M = 1$), so (3.1) does not, in general, hold. This is case of large elections, which was the focus of much of the previous literature.

Does (3.1) ever hold, and what are the combinations of parameter values under which it does? It turns out that there are certain ranges of parameter values under which sincere voting equilibrium exists. This is reflected in the following proposition:

Proposition 16. *Suppose that* $a \in \left(\frac{1-p}{p}, \frac{p}{1-p}\right)$. *Then there exist* $\varepsilon_1, \varepsilon_2, \varepsilon_3 > 0$ *such that (3.1) holds when either of the following is true:*

- $\gamma \in \left[\frac{1}{2}, \frac{1}{2} + \varepsilon_1\right]$
- $k \in \left[\frac{n}{2} - \varepsilon_2, \frac{n}{2} + \varepsilon_2\right]$
- $p \in [1 - \varepsilon_3, 1]$

Proof. See Appendix.

Proposition 16 establishes the existence of sincere voting equilibrium in several situations. First, sincere voting equilibrium exists when the minority is sufficiently similar in size to the majority - in other words, when the group is sufficiently heterogeneous in its preferences.

Second, sincere voting by all agents is also an equilibrium when the voting rule is close to a simple majority rule.

Third, when signals are very precise, sincere voting is also an equilibrium. This is an intuitive result - if signals are very precise, every voter is close to being fully informed, so she does not lose by voting sincerely.

3.4 Conclusion

This chapter has looked at voting in situations when agents have imperfect private information about a payoff-relevant state of the world. The particular aim of the chapter was to analyse conditions under which each voter behaves as if he was the only agent making the decision.

It was found that sincere voting is an equilibrium strategy for every voter under certain conditions. In particular, this happens when preferences are sufficiently heterogeneous.

A sincere voting equilibrium also exists when the voting rule is close to simple majority, or when signals are very precise.

Chapter 4

Public Opinion and Political Censorship

4.1 Introduction

Most authoritarian regimes restrict media freedom, but the degree to which they do so differs widely. Some governments, such as that of North Korea, allow very little information to reach the citizens. Others allow comparatively more freedom; as Egorov et al. (2009) observe, some dictatorships even allow greater media freedom than some European Union member states. Often, the degree of media freedom in the same regime changes over time; liberalisation of the Soviet media during perestroika is one example. The use of censorship by non-authoritarian regimes (e.g. in wartime) varies as well. For example, the United States allowed its media far more freedom when reporting from the Vietnam war than during previous wars (Hallin (1989)). Several potential explanations for the differences in the levels of media freedom have been given, including differences in censorship costs (Besley and Prat (2006)), in the amount of natural resources in the country (Egorov et al. (2009)), as well as in the size of advertising markets and the structure of government's incentives (Gehlbach and Sonin (2014)).

This chapter provides a new explanation for the degree to which a government censors its media. Namely, it examines the role of the distribution of citizens' views about the government. It thus accounts for the fact that citizens differ in their attitudes towards the regime: some citizens are very skeptical, and only willing to support the government if they see that it is doing a very good job. Others may be more supportive of the government, and willing to back it unless the news are very bad. When the government decides which news to censor and which to disclose, it takes into account the existing attitude of the public.

The news in this model are represented by a state of the world $\omega \in [0, 1]$, and the role of the government is to choose which states are concealed from the citizens, and which states are disclosed. After the government chooses its censorship strategy, the state is realised, and citizens either learn it or not. Each citizen can then take an action 0 or 1. Action 1 is favourable for the government (it can mean, for example, voting for the government in elections, or refusing to join an anti-government demonstration), and the government wants to maximise the proportion of citizens who take it. Each citizen, however, only benefits from taking action 1 if the value of the state is higher than her bias $b \in [0, 1]$. Hence, a citizen with a lower bias is more favourably predisposed towards the government.

It turns out that the distribution of biases has an important effect on the optimal censorship strategy that the government selects. A favourable change in the number of extreme voters[1] leads the government to censor a smaller set of states. On the other hand, when the *overall distribution* of biases is more favourable for the government[2], the government chooses to conceal a larger subset of states. Thus, more popular authoritarian regimes should impose stronger political censorship, while a decline in popularity induces a rational ruler to liberalise the media.

To see the intuition behind this result, consider a simplified version of the model, in which all the citizens have the same bias $\tau \in (0, 1)$ and the state can only take values 0 or 1. In that case, the government has only two pure strategies: separate the states, or pool them. Suppose that the state 1 occurs with prior probability p. If the government separates the states, it will get the votes of all the citizens with probability p; and will receive no votes with probability $1 - p$. On the other hand, if the government pools the states, then the citizens' posterior expectation of the state is p. Under pooling, the government will receive all votes with probability 1 if $\tau < p$, and will receive no votes if $\tau > p$. Hence, if τ is sufficiently low - i.e. if the citizens are biased in favour of the government - then concealing both states is best for the government. If, on the other hand, τ is high - i.e. the citizens are biased against the government - concealing the states means that the government will receive no support. Then it is better for the government to reveal the state.

This result has several implications. First, it suggests a casual link from public opinion to censorship policy. There has been much work on analysing the impact of information

[1] That is, a fall in the number of citizens with a high bias, or an increase in the number of citizens with a low bias.

[2] A more formal definition will follow.

manipulation on citizens' opinions political opinions[3]). This chapter finds that the reverse effect is also present.

Second, the fact that a more popular regime will censor a larger set of news implies several predictions. Consider, for example, the frequently debated question of whether economic growth (for example, in China) will be followed by democratisation. Traditionally[4], it has been argued that economic development typically leads to liberalisation. On the other hand, some recent observers of China - e.g. Bueno de Mesquita and Downs (2005) - have questioned this conclusion. This chapter suggests one possible channel through which strong economic performance may lead to *less* political freedom - by making citizens less inclined to oppose the government, economic growth might lead an optimising ruler to tighten censorship.

Similarly, if an outbreak of war causes the population to rally behind the government, the latter is likely to expand censorship. On the other hand, a government involved in an unpopular war will respond by providing greater media freedom. This was, indeed, the chosen action of the U.S. government during the (unpopular) war in Vietnam, compared to World War II or the Korean war, which had more support.

On the other hand, when facing a decline in popular support, a rational authoritarian ruler may choose to increase media freedom. Perestroika in the USSR, in which economic troubles coincided with the government's decision to liberalise the media, is one such example. More generally, it has been observed that a ruler is more vulnerable when he is trying to liberalise[5]. This chapter suggests one mechanism behind this effect: political liberalisation, at least in the area of media freedom, may be a government's rational reaction to a loss of public trust.

The rest of the chapter is structured as follows. Section 5.2 describes the game between the government and the public, and discusses the interpretation of the model. Section 4.3 analyses the model. It derives the necessary conditions for the shape of the censorship policy, and then describes the comparative statics for some usual shapes of the distribution of public opinion. Finally, Section 5.8 concludes. All proofs are in the Appendix.

[3] For example, DellaVigna and Kaplan (2007) and Gerber et al. (2009) look at the effect of media bias on voting in US elections. Chiang and Knight (2011) looks at the effect of newspaper endorsements. Other studies have looked at the effect of media on voting outcomes and support of the government in Peru and Brazil (Boas (2005)), Mexico (Lawson and McCann (2005)), Russia (Enikolopov et al. (2011)), Ukraine (Dyczok (2006)), and East Germany (Kern and Hainmueller (2009)

[4] See Lipset (1959).

[5] Writing about the French Revolution, Alexis de Tocqueville has famously noted that "the regime which is destroyed by a revolution is almost always an improvement on its immediate predecessor, and experience teaches that the most critical moment for bad governments is the one which witnesses their first steps toward reform" (Goldhammer and Elster (2011)).

4.1.1 Related Literature

The chapter is related to several strands of literature. First, a number of papers have looked at the determinants of government censorship policy. In Egorov et al. (2009), a ruler who chooses whether to conceal bad news faces a tradeoff between decreasing the probability of revolt, and decreasing his own ability to monitor bureaucrats. In the presence of resource rents, the latter incentive is weaker, and hence the media is made less free. Edmond (2013) looks at a coordination game in which citizens can act collectively to overthrow the regime when it is sufficiently weak; the regime, in turn, manipulates information to convince the citizens of its strength. Gehlbach and Sonin (2014) examine a government that is interested in encouraging citizens to act in a particular way (for example, to vote for it), but also in earning advertising revenue from the media outlets it owns. As in this chapter, citizens are aware of media censorship when it is present, and thus become less willing to use government-controlled media as a source of news. Shadmehr and Bernhardt (2015) look at the ability of a government to commit to media censorship. They find that a ruler gains from being able to commit. Guriev and Treisman (2015) model a dictator that can use censorship (among other instruments) to prevent an informed elite from communicating information to an uninformed majority of citizens.

In all of these papers, citizens are ex ante homogeneous in terms of their preferences for supporting or opposing the government[6]. In contrast, this chapter models a heterogeneous society and shows how the distribution of citizens' preferences affect the equilibrium choice of censorship policy.

This chapter also relates to studies of media bias (Baron (2006); Gentzkow and Shapiro (2006); Bernhardt et al. (2008); Duggan and Martinelli (2011); Anderson and McLaren (2012)). These papers look at the structure of media markets and competition between media outlets, as well as on the choice of news sources by consumers. They do not focus on the government as a strategic agent affecting the availability of news.

From a modeling point of view, the chapter is related to Rayo and Segal (2010) and Kamenica and Gentzkow (2011) who examine a more general model of information disclosure[7]. In these papers, a sender commits to an information disclosure strategy in order to persuade a receiver to take a particular action. This chapter applies more restrictive assumptions by restricting the sender to a binary action (disclose or hide) in each state, and by assuming that receiver's utility is linear in the state. On the other hand, it

[6] In Edmond (2013), citizens have identical preferences, but they can differ in the sense that they receive different signals about the underlying state. But these signals are iid, so citizens are homogeneous ex ante.

[7] See also Kolotilin (2014) for a related work.

models a situation in which there is a continuum of receivers with different utilities[8], and focuses on the comparative statics with respect to changing the distribution of the receiver's preferences.

Somewhat less closely related are papers of verifiable communication in the absence of commitment (Grossman (1981); Milgrom (1981); Seidmann and Winter (1997); Koessler (2003); Mathis (2008)). The usual result in these models is that the uninformed party always learns the true state of the world[9]. In contrast, in this chapter, the government commits to a censorship policy before the state of the world is realised. This makes it possible for full disclosure not to be optimal and for censorship to exist in equilibrium - as, indeed, is often observed in political settings.

4.2 Model

There are two groups of players: a government, and a continuum of citizens with mass 1. Each citizen i has a bias $b_i \in [0, 1]$, which represents her political position. Lower b_i means that the citizen is ex ante more supportive of the government. Biases are distributed according to a distribution F (with the associated continuously differentiable density f), which describes the distribution of public opinion.

There is a state of the world $\omega \in [0, 1]$, representing the news that the government chooses whether to disclose. Higher ω represents better news. Nature draws ω from an atomless distribution G, with density g that is strictly positive everywhere on $[0, 1]$.

For each realisation of the state, the government can choose whether to disclose it or hide (censor) it. The government's censorship strategy can thus be represented by a set $S \subseteq [0, 1]$ of states that it chooses to hide. I will assume that S is either an empty set (this corresponds to the government disclosing all the news), or a finite union of disjoint closed intervals. Thus, either $S = \emptyset$ or $S = \bigcup\limits_{j=1}^{n} [p_j, q_j]$, where $0 \leq p_j < q_j < p_{j+1} \leq 1$, $\forall j = 1, ..n$.

A citizen can take an action $a_i \in \{0, 1\}$. Let $a_i(b)$ denote an action taken by a citizen with bias b.

[8]Rayo and Segal (2010) allow the receiver to have private information about his preference parameter (which is not unlike my setting with a continuum of receivers), but assume that the preference parameter is uniformly distributed. For a generic distribution of the receiver's preferences, only some general results can be derived. On the other hand, under the more restrictive assumptions of this chapter, the effects of changes in the distribution of citizens' preferences can be analysed.

[9]Exceptions to the full disclosure result have largely been due to uncertainty over whether the informed agent has precise information (Shin (1994)), or due to informed agent's preferences being either uncertain (Wolinsky (2003)) or non-monotone in decision-maker's action (Giovannoni and Seidmann (2007)).

The government's payoff equals $\int_0^1 a_i(b)\, dF(b)$. Thus, the government wants to maximise the share of the population that takes action 1. We can think of action 1 as, for example, voting for the government in an election, taking part in a pro-government demonstration, refusing to take part in an anti-government demonstration, volunteering for military service, donating money to some government-run campaign, etc. I will refer to taking action $a_i = 1$ as *supporting the government.*

A citizen who does not support the government receives a payoff of zero. A citizen who supports the government receives a payoff of $\omega - b_i$. Thus, a citizen wants to support the government if and only if the state is high enough (that is, if the news are good enough). Bias b_i reflects how good the news should be for the citizen to be willing to choose $a_i = 1$. If b_i is close to zero, i is willing to support the government even if it is doing very poorly; while if b_i is close to one, i will oppose the government unless the news is very good.

The timing of the game is as follows. First, the government commits to a disclosure policy by choosing S, which the citizen learns. Then, Nature draws ω from G. Next, if the state ω falls in S, citizens learn it; otherwise, they update their beliefs. They then simultaneously choose actions $a_i \in \{0, 1\}$. Finally, payoffs are realised.

4.2.1 Discussion of the Model

One feature of the model is that each citizen has a bias b_i that is independent of the news. This makes it possible to separate the effect of public opinion on the equilibrium censorship policy from the reverse effect. We can think of the bias as a reflection of the citizen's attitude towards the government due to factors other than the state ω. For instance, if ω relates to how well the government is managing the economy, then b can show a citizen's level of support for its foreign policy.

Alternatively, this setup can be applied to a political competition model in which ω as represents the valence of the incumbent candidate relative to that of the challenger. The incumbent, having a degree of control over information flow, can manipulate information to convince voters that her valence is high. On the other hand, bias b can represent how close a voter is the incumbent's and challenger's chosen positions, with higher b indicating greater support for the challenger's position[10].

Additionally, the model assumes that the Sender can hide the news or reveal it up to a subset of the news space, but he cannot lie. In many situations, this is straightforward.

[10] I thank Miguel Ballester for suggesting this interpretation.

For instance, a government can either allow independent media to access certain information[11], or it can deny such access - but it cannot compel independent media to lie. Another interpretation is that, if the game is repeated indefinitely, reputational losses can be so large that lying is never an optimal strategy. Finally, we may think of revealing information as providing hard evidence of the news (for instance, photos, videos, or testimony by independent media) - evidence which cannot be falsified. Under this interpretation, lying (or telling the truth with no evidence to support it) leads the citizens to ignore the government's message, and is thus identical to concealing information.

Note that the government commits to a censorship strategy S before the state is realised. This setup reflects a number of real-world situations. Consider, for example, restrictions on hate speech or incitement of violence, which are in place in many countries. These restrictions are typically specified in laws that are approved in advance and cannot be changed even when the government feels it is advantageous to release a particular piece of information[12]. Even in countries without rule of law, censorship of politically sensitive news is often regulated by bureaucratic instructions, rather than by decisions that are made every time the news arrive[13]. Another form of commitment is establishing constitutional provisions for a certain degree of freedom of speech[14]. Commitment may also arise as a credible equilibrium strategy in a repeated interaction.

Another feature of the model is that citizens know the government's censorship policy, and, if the state is not revealed to them, form their beliefs about the state based on that knowledge. Chiang and Knight (2011) provide some empirical evidence that voters are indeed sufficiently sophisticated to be able to detect media bias and to interpret media reports accordingly. Furthermore, if a particular censorship policy has been in place for some period of time, citizens may be particularly able to learn what kind of news are not revealed to them[15].

[11] For instance, by releasing classified documents. Similarly, a government engaged in a war can allow independent journalists to travel to the warzone.

[12] We do not normally think of hate speech or incitement of violence as conveying information. However, we can imagine that some types of e.g. racist incitement may be convincing to the public, while others serve only to show the speaker and his racist message in a bad light. A planner that wants to reduce the amount of racism in a society may want to block convincing messages but allow those that discredit the speaker. In this sense, an effective racist demagogue is "bad news", while an ineffective one is "good news".

[13] A study of press censorship in 19th century Europe by Goldstein (2000) mentions a large number of censorship laws and bureaucratic circulars issued to newspapers by various governments. Kris (1941) describes a twenty-page set of instructions, given to Czechoslovak newspaper editors by the Nazi German occupation authorities in 1939, explaining which kinds of news stories would be allowed to be published in future. In either case, the authorities had to commit to a specific set of instructions, rather than examining every article that the newspapers wanted to publish - probably because the latter approach would be too time-consuming.

[14] Even in dictatorships, such commitment may be credible since the dictator may face international sanctions, loss of foreign aid, etc. if he breaks the commitment.

[15] A striking example of this is a 1982 accident in a Moscow Metro station, which claimed eight lives. The accident itself could not be concealed, but the Soviet media made no mention of the deaths, in accordance with the long-standing policy of not revealing such news. Most people, however, knew that

The payoff of a citizen who supports the government only depends on whether the state is higher than her bias. A citizen thus only cares about her action, and not about the number of citizen who take a particular action. We may think that a given citizen cannot affect the latter, since she is rarely pivotal. On the other hand, she may care about making a "correct" decision - supporting the government when it is doing a good job, and not supporting it when it is not. This is in line with models of expressive voting[16].

Finally, note that the government's payoff is monotone and continuous in the share of the population that supports it. On the other hand, in models of two-party political competition the government typically cares about the behaviour of the median voter, and its payoff is often discontinuous at the 50% support level. Several reasons why the payoff maybe continuous come to mind. In an autocracy, the explanation is fairly straightforward: the government is aiming to minimise the number of people that may take part in anti-government protests, or to maximise the number of people attending pro-government demonstrations. In a political system with contested elections, it may be the case that it each citizen may randomly decide not to come to the polls, or to vote for or against the government based on some random factor other than his bias and the state ω. In that case the government's payoff is continuous in the number of voters who support it.

4.3 Analysis and Results

4.3.1 General Results

Suppose the government has chosen a censorship policy S. Then if $\omega \notin S$, citizens learns the state. Citizen i's payoff from taking action 1, given her bias b_i, equals $\omega - b_i$. She thus strictly prefers $a_i = 1$ if $b_i < \omega$, and $a_i = 0$ if $b_i > \omega$.[17] The government's payoff then equals $F(\omega)$.

If S is nonempty and $\omega \in S$, citizen i's payoff from taking action 1 equals $\mathrm{E}[\omega - b_i \mid \omega \in S]$. Thus, citizen i supports the government if and only if $b_i < \mathrm{E}[\omega \mid \omega \in S]$. The government's payoff in this case equals $F(\mathrm{E}[\omega \mid \omega \in S])$.

deaths would not be reported regardless of how many are killed. As a result, many of them formed a belief that the number of causalities was much larger, and rumours began to circulate that hundreds of people had been killed. See Zolotov (2002) for the description of the accident, its reporting, and subsequent rumours.

[16] See Brennan and Hamlin (1998), Kamenica and Brad (2014).

[17] The share of citizens for whom $b_i = \omega$ is zero, so they can be ignored.

Then, if the government has chosen a non-empty set S, its expected payoff equals

$$v(S) = \Pr(\omega \notin S) \, \mathrm{E}[F(\omega) \mid \omega \notin S] + \Pr(\omega \in S) \, F(\mathrm{E}[\omega \mid \omega \in S])$$

or

$$v(S) = \int\limits_{\omega \notin S} F(\omega) \, dG(\omega) + \mu_S F(t_S)$$

where $\mu_S \equiv \int\limits_{\omega \in S} dG(\omega)$ is the measure of S, and $t_S \equiv \mathrm{E}[\omega \mid \omega \in S] = \frac{1}{\mu_S} \int\limits_{\omega \in S} \omega dG(\omega)$.

If the government has chosen $S = \emptyset$, then $\Pr(\omega \in S) = 0$, and

$$v(S) = \int_0^1 F(\omega) \, dG(\omega)$$

The government then chooses censorship policy S to maximise $v(S)$. Clearly, the optimal S depends on the shapes of F and G. In order to characterise it, we can look at the following deviation from S: take an interval $[w, w+\varepsilon] \notin S$ of states that are revealed, and hide them - i.e. pool them with S. Then if S is optimal, the marginal effect of such a change as $\varepsilon \to 0$ must be negative. Similarly, for an interval $[w, w+\varepsilon] \in S$, the marginal effect of separating these states from S (i.e. of revealing them) as $\varepsilon \to 0$ must be negative.

Formally, for a given S, we can define a function $z_S(\omega) \equiv \int_{t_S}^{\omega} f(t_S) \, dx - \int_{t_S}^{\omega} f(x) \, dx$. Intuitively, the first integral is related to the marginal effect of pooling a state ω with S, while the second integral is related to the marginal effect of separating ω from S. Then, when S is optimal, $z_S(\omega)$ must be positive for any $\omega \in S$, and negative for any $\omega \notin S$. This provides a necessary condition for S to constitute an optimal censorship policy; this condition is formally described in the following proposition:

Proposition 17. *Suppose that $\mathcal{S}$ maximises $v(\cdot)$. Then the following must hold:*

- $z_S(\omega) \geq 0$ *for any* $\omega \in S$
- $z_S(\omega) \leq 0$ *for any* $\omega \notin S$

where $z_S(\omega) \equiv \int_{t_S}^{\omega} f(t_S) \, dx - \int_{t_S}^{\omega} f(x) \, dx$.

Note that $z_S(\omega)$ can be written as $f(t_S)(\omega - t_S) - F(\omega) + F(t_S)$. Then when S is optimal, we must have $f(t_S)(\omega - t_S) \geq F(\omega) - F(t_S)$ for all $\omega \in S$, and $f(t_S)(\omega - t_S) \leq F(\omega) - F(t_S)$ for all $\omega \notin S$. Geometrically, $f(t_S)$ is the slope of the bias distribution F at t_S. Then from the condition above it follows that states ω for which $F(\omega)$ are above

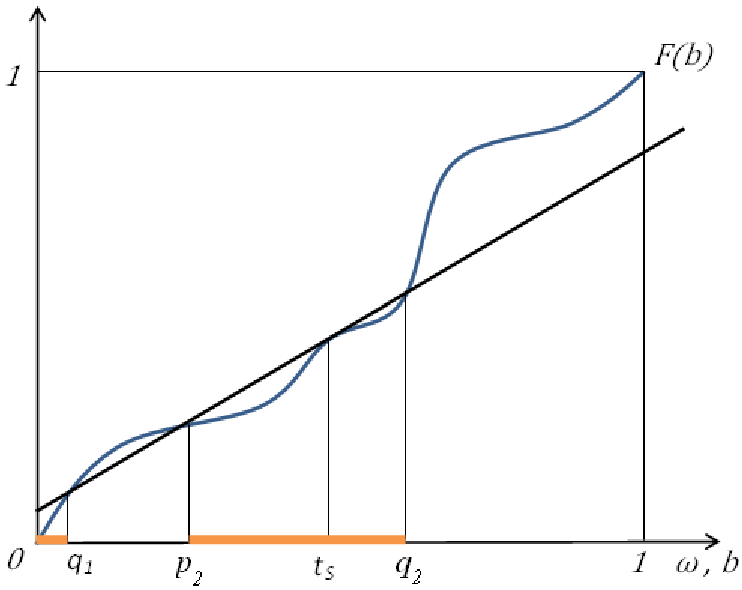

FIGURE 4.1: Illustration of Proposition 17. Here, $S = [0, q_1] \bigcup [p_2, q_2]$. The diagonal line is tangent to the bias distribution F at t_S. States ω for which $F(\omega)$ is above the tangent line are disclosed; states for which it is below the tangent line are hidden. The boundaries of S are points at which the tangent line intersects F.

the line that is tangent to F at t_S are disclosed; those for which $F(\omega)$ is below that line are not disclosed and belong to S. The boundaries of S are points ω at which the tangent line intersects F - i.e. in which $f(t_S) = \frac{F(\omega)-F(t_S)}{\omega - t_S}$. This is illustrated in Figure 4.1.[18]

From this result and from Figure 4.1, it can be immediately seen that when F is concave, it is optimal not to reveal any information, i.e. to set $S = [0, 1]$. Indeed, if F is concave, then for any S and the associated tangent line, $F(\omega)$ will be below that tangent line for all $\omega \in [0, 1]$. Similarly, if F is convex, then full disclosure - i.e. $S = \emptyset$ - is optimal, since in this case for every $\omega \in [0, 1]$, $F(\omega)$ will be above the tangent line[19]. It is possible, however, to go further and show that convexity of F is not only a sufficient, but also a necessary condition for full disclosure to be optimal. Full disclosure is optimal only if F is convex. Thus, a government faced faced with a convex distribution of biases will, at the optimum, conceal nothing from the public. Recall that a citizen whose bias is higher is less interested in supporting the government. Informally, when F is convex (and hence f is increasing), higher biases are more likely to occur. Thus, full transparency is optimal if among the citizens, unfavourable attitudes towards the government are more likely than favourable attitudes. This is similar in spirit to the more general results derived in Section 4.3.2 below.

[18]Technically, the condition $f(t_S) = \frac{F(\omega)-F(t_S)}{\omega - t_S}$ is only defined for $\omega \neq t_S$. It is also possibly for $\omega = t_S$ to be a boundary of S.

[19]These results are similar to the results obtained in Kamenica and Gentzkow (2011) on when a Sender benefits from persuasion.

Before we go to these specific results, we can note another general result about the optimal censorship policy S. As described above, S can theoretically consist of many disjoint intervals. However, this probably rarely occurs in reality - governments are unlikely to censor many distinct ranges of news. The following proposition, the proof of which is in the Appendix, shows that censorship policy S will usually consist of a small number of intervals:

Proposition 18. *If f has $m < \infty$ local weak maxima, then at the equilibrium, $n \leq m$.*

Hence, the number of intervals in the set of censored news cannot be larger than the number of local maxima in the density of biases. We can say that S is "complex" if news are censored over a large number of disjoint intervals. We can also describe F as having a "complex shape" if the density f has many peaks. Then Proposition 18 says that S cannot be more complex than the shape of F.

4.3.2 Optimal Disclosure Policies

In this section I look at optimal disclosure rules under certain classes of parameter distributions.

Two settings are particularly interesting - the one in which f, the density of bias, is unimodal, and the one in which it is U-shaped. In the former case, the cdf F of biased is convex on some interval $(0, k)$ and concave on $(k, 1)$. In the latter case, F is concave on$(0, k)$ and convex on $(k, 1)$. The former case corresponds, for example, to a society in which most individuals tend to be moderate with respect to their attitude towards the government. The latter case describes a polarised society with a large number of committed supporters and opponents of the government, and not many moderates.

Proposition 19. *The following holds for simple shapes of F:*

1. *If F strictly convex on $(0, k)$ and strictly concave on $(k, 1)$ for some $k \in (0, 1)$, then there is a unique equilibrium censorship strategy $S = [p, 1]$, characterised by the condition*
$$f\left(t_{[p,1]}\right) = \frac{F(p) - F\left(t_{[p,1]}\right)}{p - t_{[p,1]}}$$
such that $0 \leq p < k$, and $t_{p,1} > k$.

2. *If F strictly concave on $(0, k)$ and strictly convex on $(k, 1)$ for some $k \in (0, 1)$, then there is a unique equilibrium censorship strategy $S = [0, q]$, characterised by the condition*
$$f\left(t_{[0,q]}\right) = \frac{F(q) - F\left(t_{[0,q]}\right)}{q - t_{[0,q]}}$$

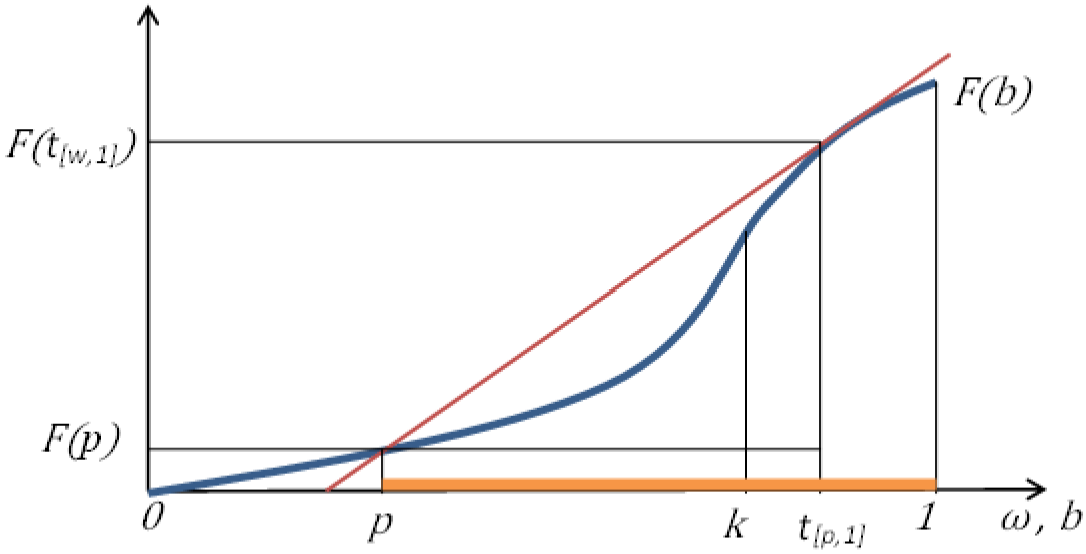

FIGURE 4.2: Illustration of Proposition 19, part 1. Here, $S = [p, 1]$ and $f\left(t_{[p,1]}\right) = \frac{F(p) - F\left(t_{[p,1]}\right)}{p - t_{[p,1]}}$.

such that $k \leq q < 1$*, and* $t_{[0,q]} < k$*.*

Hence, the unimodal and the U-shaped distributions induce disclosure policies under which either the best news, or the worst news are not disclosed. It is easy to see that full disclosure can be seen as a special case of a unimodal f with $k = 1$, or of a U-shaped f when $k = 0$. Similarly, full pooling emerges when f is unimodal and $k = 0$; or when f is U-shaped and $k = 1$.

Proposition 19 is illustrated in Figures4.2 and 4.3.

Thus, a government facing a unimodal distribution of public opinion will not provide much information to the public, except when the news are really bad. In that case, the government will disclose the news. An example of this would be a policy under which the government provides little meaningful information to its citizens, but does inform them of major natural disasters or large military defeats.

The opposite case is that of a regime which faces a polarised society that is split into hardline supporters and radical opponents of the government, with not many citizens in the middle. In that case, the government will censor all the worst news (up to a certain cutoff q), and reveal the good news.

4.3.3 Comparative Statics

It this section, we will look at the effect of possible changes in the shape of F.

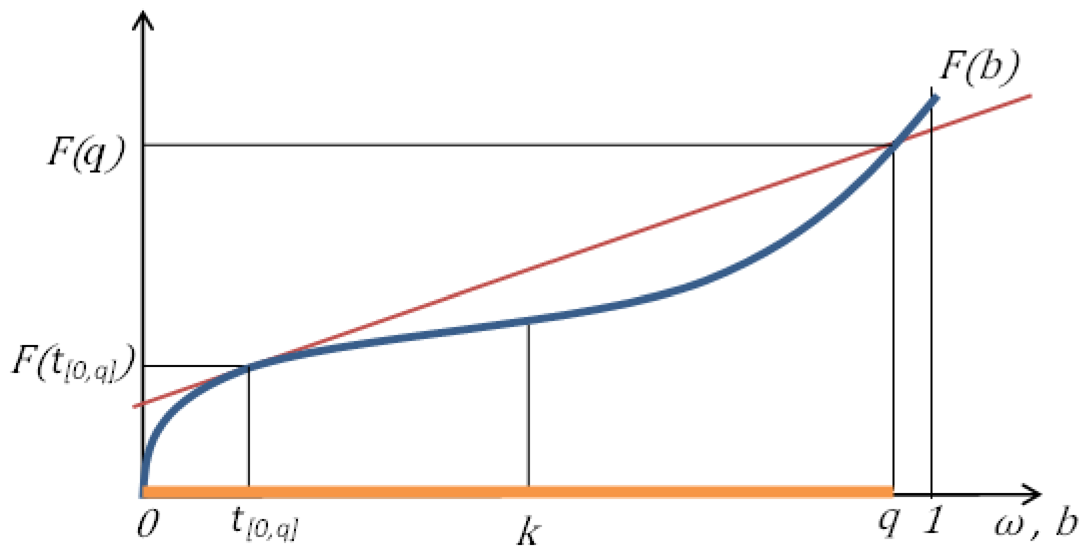

FIGURE 4.3: Illustration of Proposition 19, part 2. Here, $S = [0, q]$ and $f\left(t_{[0,q]}\right) = \frac{F(q)-F\left(t_{[0,q]}\right)}{q-t_{[0,q]}}$.

First, note that according to Proposition 19, when F is unimodal or U-shaped, the optimal disclosure policy S is determined solely by the slope of F at t_S and by the value of F at, respectively, p and q. Thus, any change in F that preserves its general shape (unimodal or U-shaped) and the abovementioned parameters will leave the optimal censorship policy unchanged.

Suppose, on the other hand, that a uniform F with peak k is replaced with another uniform distribution $\hat{F}$ whose peak is $\hat{k} > k$. Moreover, suppose that $\hat{k}$ is sufficiently far to the right that $\hat{k} > t_{[p,1]}$. Since the distribution remains uniform, the new censorship strategy should be $\hat{S} = [\hat{p}, 1]$. From Proposition 19 we know that $t_{[\hat{p},1]}$ must be larger than $\hat{k}$. Thus, $t_{[\hat{p},1]} > t_{[p,1]}$. Therefore, $\hat{p} > p$, and $\hat{S} \subset S$ - the set of censored news is smaller.

Similarly, suppose that a U-shaped F that has a lowest point at k is replaced with another U-shaped distribution $\tilde{F}$ with a lowest point $\tilde{k} < k$. Furthermore, suppose that the change is "large enough" that $\tilde{k} < t_{[0,q]}$. We know that the new censorship strategy should be $\tilde{S} = [0, \tilde{q}]$. From Proposition 19 we know that $t_{[0,\tilde{q}]}$ must be smaller that $\tilde{k}$. Thus, $t_{[0,\tilde{q}]} < t_{[0,q]}$. Therefore, $\tilde{q} < q$, and $\tilde{S} \subset S$. Again, the set of censored states is smaller.

These results are summarised in the following proposition:

Proposition 20. *Take a distribution F that induces a censorship policy S.*

1. *Suppose that F is unimodal with peak k. Take a different unimodal distribution $\hat{F}$ with peak $\hat{k}$ such that $\hat{k} > t_S$. Then $\hat{F}$ induces a censorship policy $\hat{S}$ such that $\hat{S} \subset S$.*

2. *Suppose that F is U-shaped with the lowest point k. Take a different U-shaped distribution $\tilde{F}$ with the lowest point $\tilde{k} < t_S$. Then $\tilde{F}$ induces a censorship policy $\tilde{S}$ such that $\tilde{S} \subset S$.*

To interpret Proposition 20, recall that a larger bias b means that the citizen is less willing to support the government. For a unimodal distribution F, a shift of the peak to the right means that the modal citizen becomes more predisposed against the government. Similarly, for a U-shaped F, a shift of k to the left also means that the distribution is, in a sense, less favourable for the government. In both cases, an optimising government will react by censoring a smaller set of news.

Thus, Proposition 20 derives a partial order on the set of possible distributions of public opinion that implies that a regime that faces less favourably-minded public will, at the optimum, pursue a more liberal media policy, censoring a smaller set of news. On the other hand, a popular regime will have a more restrictive censorship policy.

Several further results can be derived. Suppose that a unimodal distribution F is replaced with another unimodal distribution $\hat{F}$ for which $\hat{f}\left(t_{[p,1]}\right) = f\left(t_{[p,1]}\right)$ but $\hat{F}(p) < F(p)$. This means that the extreme left tail of the density $\hat{f}$ is thinner than that of the density f. Now $\hat{F}(p)$ would lie below the tangent line, and the government strictly prefers to hide states in a neighbourhood of p. Hence, a smaller share of citizens with a very low bias is associated with more censorship at the optimum.

Similarly, consider a U-shaped distribution F, and replace it with another U-shaped distribution $\tilde{F}$ for which $\tilde{f}\left(t_{[0,q]}\right) = f\left(t_{[0,q]}\right)$ but $\tilde{F}(q) < F(q)$. This means that the extreme right tail of the density $\tilde{f}$ is thicker than that of the density f. Now $\tilde{F}(q)$ would lie below the tangent line, so the government strictly prefers to hide states in the neighbourhood of q. Thus, a larger share of citizens with a very high bias means that at the equilibrium, the government censors more news.

The following summarises these results:

Proposition 21. *Take a distribution F that induces a censorship policy S.*

1. *Suppose that F is unimodal. Take a different unimodal distribution $\hat{F}$ such that $\hat{f}\left(t_{[p,1]}\right) = f\left(t_{[p,1]}\right)$ and $\hat{F}(p) < F(p)$. Then $\hat{F}$ induces a censorship policy $\hat{S}$ such that $S \subset \hat{S}$*

2. *Suppose that F is U-shaped. Take a different U-shaped distribution $\tilde{F}$ such that $\tilde{f}\left(t_{[0,q]}\right) = f\left(t_{[0,q]}\right)$ and $\tilde{F}(q) < F(q)$. Then $\tilde{F}$ induces a censorship policy $\tilde{S}$ such that $S \subset \tilde{S}$*

To interpret Proposition 21, recall that a thinner left tail of f means that the number of extreme pro-government voters is smaller. At the same time, a thicker right tail of f implies a larger number of extreme anti-government citizens. Both changes result in the government censoring a larger subset of states at the equilibrium. Thus, while Proposition 20 shows that a shift of the "mainstream" opinions (in the unimodal case, a shift of the modal citizen) away from supporting the government leads the latter to relax censorship restrictions, Proposition 21 suggests that an unfavourable change in the number of "extreme" voters causes the government to censor a larger set of news.

We can also use Proposition 21 to examine the effect of radicalisation - that is, of a shift of the mass of biases b towards the tails. When the society is dominated by moderates - as in the case of a unimodal distribution - a shift of public opinion away from the centre results, according to the above logic, in liberalisation. On the other hand, when the society is already radicalised (which corresponds to the U-shaped F), further radicalisation increases the equilibrium level of censorship.

4.4 Conclusions

This chapter has examined the link between political censorship and public opinion. There has been much research on the effect of information manipulation on the views of the public. This chapter, on the other hand, has analysed the reverse effect: the impact of public opinion on the government's censorship policy.

It has been found that a government that faces more favourable public opinion overall will censor a larger set of news. This suggests that popular authoritarian regimes are likely to impose stricter censorship, while a loss (gain) of popularity can lead a regime to increase (reduce) media freedom.

On the other hand, an favourable change in the number of radical supporters or opponents can lead a government to strengthen censorship. Radicalisation in a largely moderate society causes the regime to increase media freedom, while radicalisation in an already polarised society leads to greater censorship.

The result that more popular governments are likely to censor more news leads to fairly pessimistic predictions about the likely political effects of economic growth in autocracies. Good economic performance, as well as other developments that make citizens more satisfied with the regime, will lead the rulers to reduce media freedom. Liberalisation is likely to happen as a result of economic shocks or other changes that reduce the government's popularity.

Chapter 5

Competition and Information Disclosure

5.1 Introduction

Consider a group of applicants competing for a place at a university. The university would like to admit the applicant with the highest ability, but the ability of each applicant is his private information. Applicants, however, can credibly communicate their ability, at a cost. For example, each of them can take a costly test, increase her exposure by participating in conferences and other events, or simply spend more time writing her application essay or cover letter. If the application process becomes more competitive, will the applicants be more likely or less likely to communicate their ability?

More formally, suppose that a number of contestants are competing for a prize. Each contestant has a type, which is her private information. The prize is allocated by a decision-maker, who would like to give it to the contestant with the highest type. Each contestant can decide to disclose her type. Disclosure is credible and costly. Contestants decide whether to reveal their types simultaneously.

The paper asks several questions. First, is information more likely or less likely to be communicated when the number of contestants increases? Second, how is communication affected when the test becomes noisy? Third, will the decision-maker ever want to make the test compulsory, and if yes, when?

While competition for university admission is one straightforward application of this model, the setting is relevant to a number of other situations. For example, workers competing for a job can choose to signal their skill by acquiring education at some cost. Candidates for political office can communicate their competence to voters by running a costly campaign. A firm wishing to credibly reveal the quality of its product can submit it for certification by independent experts.

The model produces several novel results. First, the effect of competition on information revelation depends on the cost of revealing the type. When that cost is very high, contestants never disclose their type. When it is moderately high, increasing the number of contestants increases the amount of information being disclosed. But when the cost is below a certain threshold, increasing the number of contestants makes it *less* likely that they take the test.

Second, as the number of contestants goes to infinity, the probability that a given contestant reveals her type approaches zero. But at the same time, the probability that *some* contestant discloses her type approaches a constant that is distinct from zero and from one. Hence, even in the limit, there are always some contestants who disclose their type, and others who do not.

Third, I consider what happens to the equilibrium when the test reveals a contestant's type with some noise. It turns out that contestants are always more likely to take a costly test if it reveals their type imperfectly. In general, greater noise makes it more likely that a contestant takes the test.

Finally, I consider what happens if the decision-maker has an option to make the test compulsory. It turns out that as long as the number of contestants is above a certain threshold, making the test compulsory is never optimal for the decision-maker. That threshold need not be very high: for example, when the contestants' types are uniformly distributed, compulsory testing is suboptimal as long as the number of contestants is greater than two.

The rest of the paper is structured as follows. Section 5.2 describes the baseline model, which is analysed and solved in Section 5.3. Section 5.4 discusses the welfare effects of a change in the number of contestants. Section 5.5 shows that as long as the number of contestants is sufficiently large, making the test compulsory is not an optimal strategy for the decision-maker. Section 5.6 shows that a noisier test is always more atractive for the contestants. Section 5.7 extends the analysis to the case when the cost of taking the test may depend on the type. Finally, Section 5.5 concludes.

5.1.1 Related literature

Several papers have looked at information disclosure in competitive settings. Gentzkow and Kamenica (2011) (Gentzkow and Kamenica (2011), Gentzkow and Kamenica (2015)) examine a setting in which senders can communicate information to a receiver about an unknown state of the world. They do so by commiting to generic experiments. Bhattacharya and Mukherjee (2013) develop a model in which experts with different agenda can reveal a state of the world to a decision-maker. In these models, information disclosure is costless, and competition is primarily interpreted as the degree to which the agents' prefrences are aligned.

The industrial organisation literature has looked at information disclosure by competing firms. Board (2009) develops a model of two firms, each of whom chooses whether to disclose the quality of its product. In a similar setting, Levin et al. (2005) compares information disclosure by a cartel to a duopoly situation. These papers do not examine the effect of an increase in the number of firms. Forand (2013) compares information provision under monopoly to a duopoly case. In that paper, information provision is costless. Stivers (2004), and Ivanov (2013) also look at the effect of competition on information revelation when the latter is costless. They find that greater competition tends to imply greater information disclosure. Cheong and Kim (2004) look at the effect of firms choosing whether to reveal information about their product quality in a competitive market. In their paper, not revealing the quality means that the firm receives zero profit, and if quality is uniformly distributed, an increase in competition results in a decrease in the amount of information being revealed, regardless of the cost. When the number of firms approaches infinity, no firm discloses information, unlike in this paper. Janssen and Roy (2015) analyse a duopoly in which firms choose between disclosing their product quality at some cost, and signalling quality through a choice of price. They find that even when disclosure cost is arbitrarily small, non-disclosure is an equilibrium.

This paper is also related to the literature on auctions with costly participation (Stegeman (1996); Menezes and Monteiro (2000); Moreno and Wooders (2011)). In these models, each buyer can choose to either take part in an auction, paying the cost in exchange for a chance to win; or to abstain from participating. Conversely, a contestant who does not take the test still has a positive probability of winning the prize when nobody else takes the test. Because of this, it is possible that the probability of taking the test decreases if competition becomes larger.

5.2 Model

There are $n > 1$ contestants (male) that are competing for a prize which is allocated by a decision-maker (female). The value of the prize to each contestant is 1. Each contestant i has a type $x_i \in [0, 1]$, which is her private information. Types are drawn independently from a distribution F with full support on $[0, 1]$. The decision-maker receives a payoff x_i if she allocates the prize to contestant i.

Before the decision-maker allocates the prize, each contestant decides whether to take a test. Taking the test perfectly reveals the type to the decision-maker[1]. Contestant i's cost of taking the test equals $c \in (0,1]$. Thus, in the basic model, the cost of taking the test does not depend on the type[2].

The decision-maker would like to allocate the prize to a contestant with the highest type. If the decision-maker's posterior belief is such that several contestants have the highest type, she randomises between them uniformly. One way of interpreting this is to say that with an equal probability, each contestant has a quality that makes the decision-maker lexicographically prefer him to other contestants.

The timing is as follows. First, nature draws x_i for every contestant i. Each contestant learns his type. Then, contestants simultaneously decide whether to take the test. The decision-maker then learns the types of those contestants who take the test. Finally, the decision-maker chooses a contestant that receives the prize.

5.3 Communication

The strategy of every contestant i is a function $h_i : [0,1] \to [0,1]$ which maps the contestant's type to the probability of taking the test. The paper will focus on symmetric strategies of the kind $h_1 = ... = h_n \equiv h$.

The decision-maker will allocate the prize to a contestant whose ex post expected type is the highest. Intuitively, at the equilibrium, a contestant whose type is close to zero should not take the test, because he is unlikely to win. If the type is higher, the probability of winning is (weakly) larger. This logic implies the following proposition:

Proposition 22. *At every symmetric equilibrium, for a given n there exists a threshold $b_n \in [0,1]$ such that $h(x) = 1$ for all $x > b_n$, and* $\Pr[h(x) > 0 \mid x < b_n] = 0$.

Proof. See Appendix.

In words, a symmetric equilibrium involves a threshold $b_n \in [0,1]$ such that all types above b_n are revealed with probability 1, and all types below b_n are revealed with probability 0 - except for, possibly, some set of types whose mass is zero. This last possibility is largely irrelevant, because the paper focuses on what happens in expectation. I will

[1] Section 5.6 extends the analysis to the case when taking the test produces a noisy signal about the contestant's type.

[2] In Section 5.7, I will allow the cost to potentially increase with type. This can reflect the fact that, for example, taking the test may be harder for a university applicant with lower ability; or that signaling skill by acquiring education may be more costly for a job applicant with lower skill, in the spirit of Spence (1973).

thus assume that the equilibrium strategy of each contestant takes a simple form of never revealing any type below b_n.

The decision-maker's expected payoff equals the expected type of the contestant whom she gives the prize. At a Bayesian equilibrium, if contestant i has a type above b_n (and thus takes the test), the decision-maker learns his type with certainty. Hence, if at least one contestant takes the test, the decision-maker is able to allocate the prize to the best contestant with certainty. In these situations, I will say that the decision-maker *makes an informed decision.*

If contestant i does not take the test, the decision-maker's expectation of i's type equals $\mathrm{E}_F(x \mid x < b_n)$, where $\mathrm{E}_F(\cdot)$ denotes expectation taken over F. This expression is well-defined when $b_n > 0$. Note that since the test is costly, if $b_n = 0$ there is always an incentive to deviate when the type is sufficiently close to zero, as the probability of winning the prize after revealing the type is very low. Hence, $b_n = 0$ cannot be an equilibrium as long as $c > 0$.

Since $b_n > \mathrm{E}_F(x \mid x < b_n)$, a contestant who does not take the test has a lower ex-post expected type that any contestant who does. He can thus only win the prize if nobody else takes the test, in which case the decision-maker selects him with probability $\frac{1}{n}$. Thus, if a contestant does not take the test, his overall probability of winning the prize is $F(b_n)^{n-1}\frac{1}{n}$. On the other hand, a contestant with type $x_i > b_n$ takes the test and wins the prize if every other contestant has a lower type - which happens with probability $F(x_i)^{n-1}$.

Suppose that $c \leq \frac{n-1}{n}$. At $x_i = b_n$, contestant i must be indifferent between taking and not taking the test, which gives us the equation:

$$F(b_n)^{n-1} - c = F(b_n)^{n-1}\frac{1}{n} \tag{5.1}$$

Hence,

$$F(b_n) = \left(\frac{cn}{n-1}\right)^{\frac{1}{n-1}} \tag{5.2}$$

On the other hand, if $c > \frac{n-1}{n}$, then the left-hand side of (5.1) is smaller than the right-hand side for all $b_n \in [0,1]$. Hence, the equilibrium strategy of every contestant is to never reveal the type. In that case, the threshold b_n equals 1. Thus, the unique symmetric equilibrium is given by

$$F(b_n) = \min\left\{\left(\frac{cn}{n-1}\right)^{\frac{1}{n-1}}, 1\right\}$$

or

$$b_n = \begin{cases} F^{-1}\left[\left(\frac{cn}{n-1}\right)^{\frac{1}{n-1}}\right] & \text{if } c \leq 1 - \frac{1}{n} \\ 1 & \text{if } c > 1 - \frac{1}{n} \end{cases}$$

There are two natural ways of measuring the degree to which information is disclosed. One such indicator is $F(b_n)$. From the decision-maker's point of view, $F(b_n)$ is the probability that a given contestant does not take the test.

Another useful indicator is $F(b_n)^n$. This is the probability that there is no communication - i.e. that even the contestant with the highest type does not take the test. To put it in a different way, $1 - F(b_n)^n$ is the probability that the decision-maker makes an informed decision.

Unsurprisingly, both $F(b_n)$ and $F(b_n)^n$ weakly increase if c goes up - when the test is more costly, contestants are less likely to take it. A more interesting question is what happens if n increases.

If $c > \frac{n-1}{n}$, then $F(b_n) = F(b_n)^n = 1$, and a further increase in n does not change b_n. On the other hand, suppose that $c < \frac{n-1}{n}$. Then the values of $F(b_n)$ and of $F(b_n)^n$ change with n.

Suppose that the number of contestants changes from n to $n+1$. Then the probability that a contestant takes the test decreases whenever $F(b_n) < F(b_{n+1})$, i.e. whenever

$$\left(\frac{cn}{n-1}\right)^{\frac{1}{n-1}} < \left(\frac{c(n+1)}{n}\right)^{\frac{1}{n}}$$

this holds if and only if

$$c < \frac{n-1}{n}\left(\frac{n^2-1}{n^2}\right)^{n-1}$$

Note that $\frac{n-1}{n}\left(\frac{n^2-1}{n^2}\right)^{n-1}$ is less than $\frac{n-1}{n}$ and greater than zero. Hence, we have the following result[3]:

Proposition 23. *The probability that a given contestant takes the test decreases with n if $c \in \left[0\,,\ \frac{n-1}{n}\left(\frac{n^2-1}{n^2}\right)^{n-1}\right]$, increases with n if $c \in \left(\frac{n-1}{n}\left(\frac{n^2-1}{n^2}\right)^{n-1}\,,\ \frac{n-1}{n}\right)$, and stays constant as n changes if $c \in \left[\frac{n-1}{n}\,,\ 1\right]$.*

Similarly, we can check the effect of n on the probability that *any* contestant takes the test. This probability decreases as n goes up if and only if $F(b_n)^n < F(b_{n+1})^n$, which

[3]We can obtain a simple expression if we replace n with a continuous variable and differentiate (5.2) with respect to n. Then we can find that a marginal increase in n increases $F(b_n)$ whenever $c < \frac{n-1}{n}e^{-\frac{1}{n}}$.

holds whenever

$$\left(\frac{cn}{n-1}\right)^{\frac{n}{n-1}} < \left(\frac{c\,(n+1)}{n}\right)^{\frac{n+1}{n}}$$

or equivalently, whenever

$$c < \frac{n-1}{n}\left(\frac{n^2-1}{n^2}\right)^{n^2-1}$$

Again, note that $\frac{n-1}{n}\left(\frac{n^2-1}{n^2}\right)^{n^2-1}$ lies between zero and $\frac{n-1}{n}$. Therefore, we have the following result[4]:

Proposition 24. *The probability that any contestant takes the test decreases with n if $c \in \left[0\,,\, \frac{n-1}{n}\left(\frac{n^2-1}{n^2}\right)^{n^2-1}\right]$, increases with n if $c \in \left(\frac{n-1}{n}\left(\frac{n^2-1}{n^2}\right)^{n^2-1}\,,\, \frac{n-1}{n}\right)$, and stays constant as n changes if $c \in \left[\frac{n-1}{n}\,,\, 1\right]$.*

Note that $\frac{n-1}{n}\left(\frac{n^2-1}{n^2}\right)^{n^2-1} < \frac{n-1}{n}\left(\frac{n^2-1}{n^2}\right)^{n-1}$. Thus, when c is below $\frac{n-1}{n}\left(\frac{n^2-1}{n^2}\right)^{n^2-1}$, greater competition means that less information is transmitted - specifically, an individual contestant is less likely to take the test, and the probability that *anyone* takes the test is lower. When c is between $\frac{n-1}{n}\left(\frac{n^2-1}{n^2}\right)^{n^2-1}$ and $\frac{n-1}{n}\left(\frac{n^2-1}{n^2}\right)^{n-1}$, greater competition means that the decision-maker is more likely to make a fully informed decision, but the probability that any given contestant reveals his type is still increasing. When c is between $\frac{n-1}{n}\left(\frac{n^2-1}{n^2}\right)^{n-1}$ and $\frac{n-1}{n}$, greater increasing the number of contestants leads to a lower probability of both an individual contestant revealing his type, and of the type being revealed by anyone. Finally, when c is larger than $\frac{n-1}{n}$, the test becomes so costly that nobody takes it, and changing n has no effect.

Overall, Propostions 23 and 24 suggest that an increase of competition reduces communication when c is low, increases communication when c is moderate, and has no effect on communication when c is high.

To see the intuition behind this result, consider a contestant who is indifferent between taking and not taking the test - i.e. a contestant whose type equals b_n. For this contestant, increasing n has two effects. First, the expected payoff from taking the test goes down, because it becomes increasingly more likely that some other contestant has a higher type (i.e. $F\left(b_n\right)^{n-1}$ goes down). Second, the expected payoff from not taking the test decreases too, because, even when nobody else takes the test, the probability $\frac{1}{n}$ of being randomly selected to receive the prize becomes smaller.

But if c is low, then so if $F\left(b_n\right)$. In that case, the change in $F\left(b_n\right)^{n-1}$ is relatively large. Thus, the first effect dominates the second, and the contestant becomes less willing to

[4]If we replace n with a continuous variable and differentiate $F\left(b_n\right)^n$ with respect to n, we find that $F\left(b_n\right)^n$ increases with n whenever $c < \frac{n-1}{n}e^{-1}$.

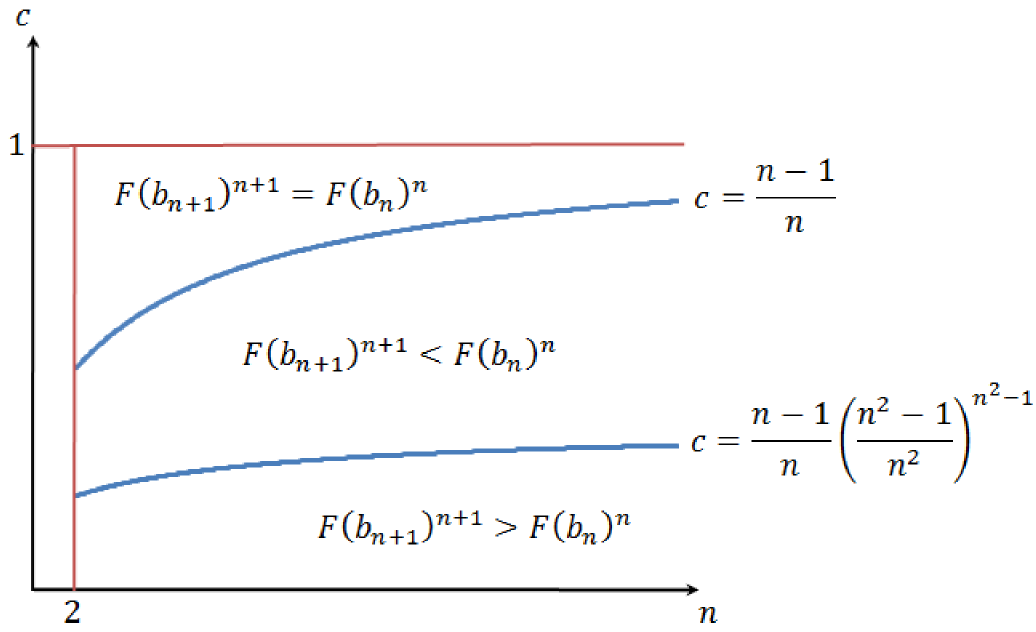

FIGURE 5.1: Effect of competition on communication.

take the test. On the other hand, if c is high, then $F(b_n)$ is close to 1. Then increasing n does not change $F(b_n)^{n-1}$ much, and so the second effect dominates the first. Finally, if c is very large, then no contestant takes the test, and a further increase in n has no effect on communication.

Figure 5.1 illustrates this result by describing the (c, n) pairs for which a change in n has these different outcomes on communication.

Another way of analysing the problem is to check what happens in the limit when the number of contestants goes to infinity. We can see that

$$\lim_{n\to\infty} F(b_n) = \lim_{n\to\infty}\left(\frac{c}{1-\frac{1}{n}}\right)^{\frac{1}{n-1}} = 1$$

Intuitively, when competition is very large, even a contestant with a high type will likely be beaten by somebody else if he takes the test. Thus, the cutoff b above which contestants take the test goes to 1.

On the other hand, if we look at the probability that *nobody* takes the test, we can see that

$$\lim_{n\to\infty} F(b_n)^n = \lim_{n\to\infty}\left(\frac{c}{1-\frac{1}{n}}\right)^{\frac{1}{1-\frac{1}{n}}} = c$$

Hence, as n becomes infinitely large, each contestant takes the test with an infinitesimal probability. Nevertheless, the probability that *anyone* (i.e. the contestant with the highest type) takes the test remains distinct from zero, and also from one. For all values of n, and also in the limit, there are (in expectation) some contestants who take the test and others who do not. Specifically, in the limit the probability that no contestant takes the test equals the cost of the test. This is summarised in the following result:

Proposition 25. *When n approaches infinity, the probability that the decision-maker makes an informed decision approaches $1 - c$.*

5.4 Welfare

Ex ante, all contestants are equally likely to win the prize. Hence, a given contestant wins the prize (and receives a payoff of 1) with probability $\frac{1}{n}$. Furthermore, at the equilibrium a contestant takes the test (and pays the cost c) whenever his type is above b. This happens with probability $1 - F(b)$. Thus, the expected utility of a contestant equals

$$u = \frac{1}{n} - [1 - F(b_n)]\, c \tag{5.3}$$

Let us look at the effect of changing the cost of the test. If $c > \frac{n-1}{n}$, then $F(b_n) = 0$, and increasing the cost of the test does not change the threshold b_n. Hence, a change in c has no effect on expected utility. On the other hand, if $c < \frac{n-1}{n}$, changing the cost changes $F(b_n)$. Using (5.2), we can see that $\frac{dF(b_n)}{dc} = -\frac{1}{n-1}\frac{F(b_n)}{c}$, and hence

$$\frac{du}{dc} = -1 + F(b_n) + \frac{dF(b_n)}{dc}c = -1 + F(b_n) + \frac{1}{n-1}\frac{F(b_n)}{c}c = -1 + \frac{n}{n-1}F(b_n)$$

Hence, $\frac{du}{dc} > 0$ if and only if $F(b_n) = \left(\frac{cn}{n-1}\right)^{\frac{1}{n-1}} > \frac{n-1}{n}$, i.e. if

$$c > \left(\frac{n-1}{n}\right)^n$$

Thus, the effect of the cost of the test on contestants' expected utility can be summarised in the following proposition:

Proposition 26. *A contestant's expected utility decreases with c if $c \in \left(0\,,\, \left[\frac{n-1}{n}\right]^n\right)$, increases with c if $c \in \left(\left[\frac{n-1}{n}\right]^n\,,\, \frac{n-1}{n}\right)$; and does not change with c if $c \in \left(0\,,\, \left[\frac{n-1}{n}\right]^n\right)$.*

Intuitively, increasing the cost of the test has two effects. On the one hand, there is the direct effect of greater cost, which reduces a contestant's utility if he ends up taking the test. On the other hand, increasing the cost reduces the probability that he ends up taking the test at the equilibrium - thus saving him the cost of the test. If the cost is low, the potential savings from not having to take the test are small, and the direct effect dominates. On the other hand, when the test is sufficiently costly, a lower probability of having to take the test has a larger impact on expected utility than the direct effect of larger cost.

5.5 Compulsory Testing

The preceding analysis has used the fact that if contestant i does not take the test, he can still receive the prize if nobody else takes the test and the decision-maker randomly assigns him the prize. This assumes, however, that the decision-maker will always allocate the prize to some contestant. Suppose, however, that the decision-maker can commit to never give the prize to a contestant who does not take the test. For example, a university can require every applicant to take the test in order for his or her application to be considered.

On the one hand, making the test a necessary condition for receiving the prize increases the incentive to take it. This reduces the threshold b_n above which a contestant takes the test. This makes it more likely that at least some contestant reveals his type. Hence, the decision-maker is more likely to be able to make an informed decision, which increases her expected payoff. On the other hand, such a commitment denies the decision-maker the chance to give the prize to any contestant in case nobody takes the test.

This section will show that if the number of contestants is sufficiently large, making the test compulsory reduces the decision-maker's expected payoff.

Consider first the case in which the test is not compulsory. With probability $1 - F(b_n)^n$, at least one contestant has a type above b_n, and takes the test. Then the decision-maker allocates the prize to a contestant with the highest type. In that case, the decision-maker's expected payoff equals

$$\mathrm{E}\left(\max\{x\} \mid \max\{x\} > b_n\right) = \frac{\int_{b_n}^{1} x d\left[F(x)^n\right]}{1 - F(b_n)^n}$$

where the above expression uses the fact that the cdf of $\max\{x\}$ is $F(x)^n$. On the other hand, with probability $F(b_n)^n$, no contestant takes the test. Then the decision-maker allocates the prize to a random contestant, and her expected payoff equals

$$\mathrm{E}\left(x \mid x < b_n\right) = \frac{\int_{0}^{b_n} x d\left[F(x)\right]}{F(b_n)}$$

Overall, the decision-maker's expected utility equals

$$v = \int_{b_n}^{1} xd\left[F(x)^n\right] + F(b_n)^{n-1} \int_0^{b_n} xd\left[F(x)\right]$$

Integrating by parts and simplifying, we obtain

$$v = 1 - \int_{b_n}^{1} F(x)^n \, dx - F(b_n)^{n-1} \int_0^{b_n} F(x) \, dx \tag{5.4}$$

Now suppose that testing is compulsory. By an argument similar to the one in Proposition 22, at a symmetric equilibrium there exists a threshold $\hat{b}_n$ such that a contestant takes the test if and only if his type is above $\hat{b}_n$. At the threshold, a contestant receives a payoff of $F\left(\hat{b}_n\right)^{n-1} - c$ if he takes the test, and zero if he does not. Indifference condition gives us

$$F\left(\hat{b}_n\right) = c^{\frac{1}{n-1}} < F(b_n)$$

The expected payoff of the decision-maker when testing is compulsory then equals

$$\hat{v} = \left[1 - F\left(\hat{b}_n\right)^n\right] \mathrm{E}\left(\max\{x\} \mid \max\{x\} > \hat{b}_n\right) = \int_{\hat{b}_n}^{1} xd\left[F(x^n)\right]$$

Integrating by parts and simplifying, we obtain

$$\hat{v} = 1 - \hat{b}_n F\left(\hat{b}_n\right)^n - \int_{\hat{b}_n}^{1} F(x)^n \, dx \tag{5.5}$$

Subtracting 5.4 from 5.5 gives us the decision-maker's expected gain from making the test compulsory:

$$\hat{v} - v = F(b_n)^{n-1} \int_0^{b_n} F(x) \, dx - \int_{\hat{b}_n}^{b_n} F(x)^n \, dx - \hat{b}_n F\left(\hat{b}_n\right)^n \equiv \Delta$$

It can be shown that this gain is negative if n is sufficiently large. In other words,

Proposition 27. *For all $c > 0$, and any F, $\lim_{n\to\infty} \Delta < 0$.*

Proof. Note that $\lim_{n\to\infty} b_n = \lim_{n\to\infty} \hat{b}_n = 1$, and $\lim_{n\to\infty} F(x)^n = 0$ for all $x \in [0,1)$. Hence, $\lim_{n\to\infty} \int_{\hat{b}_n}^{b_n} F(x)^n \, dx = 0$. Also, $\lim_{n\to\infty} F(b_n)^{n-1} = \lim_{n\to\infty} F\left(\hat{b}_n\right)^n = c$. Thus, $\lim_{n\to\infty} \Delta = c\int_0^1 F(x) \, dx - c < 0$. ∎

Thus, the results in Section 5.2 are robust to allowing the decision-maker to opt for making the test compulsory: when the number of contestants is sufficiently large, doing so is never optimal.

How large should the number of contestants be to make sure that the decision-maker will never want to make the test mandatory? When the contestants' types are uniformly distributed, it does not have to be very large, as the next result shows.

Proposition 28. *Suppose that $F(x) = x$. Then for all $c > 0$, $\Delta < 0$ whenever $n > 2$.*

Proof. See Appendix.

5.6 Noisy Signals

So far we have assumed that the test perfectly reveals the contestant's type. Suppose, however, that the test is imperfect. To be specific, suppose that rather than revealing the contestant i's type x_i, it sends a public signal $s_i = x_i + \varepsilon_i$, where ε_i is noise. For each candidate i, the nature draws ε_i from a normal distribution with mean zero and variance σ^2. Suppose that ε_i is drawn independently across i and independently of contestants' types. Candidates know their types x_i, but they do not know the noise when they decide whether to take the test.

How does the precision of the test affect the equilibrium communication? In particular, we are interested in the effect of changing σ on the probability that a contestant takes the test.

Consider a symmetric threshold strategy under which candidate i takes the test if and only if $x_i \geq b_n$, as in Section 5.3. Take a contestant i whose type equals b_n. If i does not take the test, the decision-maker knows that his type is below b_n. He can then only win the prize if nobody else takes the test, which happens with probability $F(b_n)^{n-1}$. If i takes the test, he wins the prize if every other contestant j either (i) does not take the test, or (ii) takes the test and his result s_j is lower than s_i. The probability of winning the prize after taking the test then equals

$$[F(b_n) + \Pr(x_j > b_n \cap x_j + \varepsilon_j < b_n + \varepsilon_i)]^{n-1} = [F(b_n) + \gamma]^{n-1}$$

where $\gamma \equiv \Pr(x_j > b_n \cap x_j + \varepsilon_j < b_n + \varepsilon_i) = \Pr(x_j > b_n \cap \varepsilon_j - \varepsilon_i < b_n - x_j)$.

If $F(b_n) = 1$, then nobody takes the test, and a marginal change in σ will not, in general, change this. Suppose instead that $F(b_n) < 1$. In that case, b_n is determined by the indifference condition:

$$[F(b_n) + \gamma]^{n-1} - c = F(b_n)^{n-1} \frac{1}{n} \tag{5.6}$$

This is similar to 5.1, with one difference. Namely, if b_n is held constant, then adding noise increases the payoff from taking the test for a contestant whose type equals b_n. Previously, if such a contestant took the test, he would win if and only if every other contestant had a lower type. Now he can also win if some of the other contestants have higher types, as long as their test results are worse. On the other hand, the expected payoff from not taking the test is unchanged.

Intuitively, this implies that adding noise makes taking the test a more attractive option than in the case without noise. A contestant who was indifferent when the test was not noisy (that is, a contestant whose type equals b_n) will now strictly prefer to take the test. Thus, if the test becomes noisy, b_n falls and a given contestant becomes more likely to reveal his type.

More generally, we can show that making the test noisier will increase the probability that a contestant takes the test. This is shown in the following proposition:

Proposition 29. *For any n and any c, an increase in σ weakly decreases b_n.*

Proof. See Appendix.

5.7 Type-Dependent Cost

Suppose that the cost of taking the test depends on the contestant's type. For example, in the spirit of Spence (1973) model, a job applicant who can signal his ability by obtaining education can find it easier to do so if his ability is higher.

Suppose that a contestant with type x_i can reveal his type at a cost cx_i^{-a}, with $a \geq 0$. Then the indifference condition (5.1) becomes

$$F(b_n)^{n-1} - cb_n{}^{-a} = F(b_n)^{n-1} \frac{1}{n} \tag{5.7}$$

Consider a special case in which types are drawn from a uniform distribution. Then (5.7) implies that at the equilibrium

$$F(b_n) = b_n = \min\left\{ \left(\frac{cn}{n-1} \right)^{\frac{1}{n-1+a}}, 1 \right\}$$

If $F(b_n) < 1$, then an increase in n increases $F(b_n)$ whenever

$$\left(\frac{cn}{n-1}\right)^{\frac{1}{n-1+a}} < \left(\frac{c(n+1)}{n}\right)^{\frac{1}{n+a}}$$

which holds if and only if

$$c < \frac{n-1}{n}\left(\frac{n^2-1}{n^2}\right)^{n-1+a}$$

5.8 Conclusions

This paper developed a model of costly information disclosure by contestants competing for a prize. Disclosure was modelled as a costly test, that each contestant can take to reveal his type.

Several new results were derived. First, greater competition decreases the probability that information is disclosed if the cost of disclosure is sufficiently low, and increases it if the cost of disclosure is above a certain threshold. In the limit, when the number of contestants approaches infinity, there are still some contestants who disclose information and others who do not

Second, when disclosure is noisy, the contestants become more likely to revea infformation. The noisier the test, the more likely is each contestant to reveal information.

Third, it was found that it is not optimal for the decision-maker to make the test compulsory, as long as the number of contestants is sufficiently high.

Appendix

Proof of Proposition 1

The value of ignorance for agent i is

$$d_i = (\pi_0 x_i + [1-\pi_0]\, y_i)\, g_\gamma\left(\pi_0 x + [1-\pi_0]\, y\right) - \pi_0 x_i g_\gamma\left(x\right) - (1-\pi_0)\, y_i g_\gamma\left(y\right)$$

1. If $g_\gamma(x) = g_\gamma(y) = g_\gamma(\pi_0 x + [1-\pi_0]\, y)$, then $d_i = 0, \forall i \in I$, so all agents are indifferent between learning and not learning.

2. If $g_\gamma(x) = g_\gamma(y) = 0$ and $g_\gamma(\pi_0 x + [1-\pi_0]\, y) = 1$, then $d = \pi_0 x + [1-\pi_0]\, y$, so $g_\gamma(d) = g(\pi_0 x + [1-\pi_0]\, y) = 1$.

3. If $g_\gamma(x) = g_\gamma(y) = 1$ and $g_\gamma(\pi_0 x + [1-\pi_0]\, y) = 0$, then $d = -(\pi_0 x + [1-\pi_0]\, y)$, so $g_\gamma(d) = 1 - g_\gamma(\pi_0 x + [1-\pi_0]\, y) = 1$.

4. If $g_\gamma(x) = 1$ and $g_\gamma(y) = g_\gamma(\pi_0 x + [1-\pi_0]\, y) = 0$, then $d = -\pi_0 x$, so $g_\gamma(d) = 1 - g_\gamma(\pi_o x) = 1 - g_\gamma(x) = 0$.

5. In a similar way, it can be shown that $g_\gamma(y) = 1$ and $g_\gamma(x) = g_\gamma(\pi_0 x + [1-\pi_0]\, y) = 0$ imply $g_\gamma(d) = 0$.

6. If $g_\gamma(x) = 0$ and $g_\gamma(y) = g_\gamma(\pi_0 x + [1-\pi_0]\, y) = 1$, then $d = \pi_0 x$, so $g_\gamma(d) = g_\gamma(\pi_0 x) = g_\gamma(x) = 0$.

7. In a similar way, it can be shown that $g_\gamma(y) = 0$ and $g_\gamma(x) = g_\gamma(\pi_0 x + [1-\pi_0]\, y) = 1$ imply $g_\gamma(d) = 0$.

Proof of Proposition 2

Without loss of generality, normalise the size of I to one. Recall the the committee has a collective preference for ignorance whenever $g_{0.5}(x) = g_{0.5}(y) = 0$ or $g_{0.5}(x) = g_{0.5}(y) =$

1. The former condition says that

$$\#L + \#I_X \geq \frac{1}{2} \text{ and } \#L + \#I_Y \geq \frac{1}{2} \tag{8}$$

while the latter says that

$$\#W + \#I_X \geq \frac{1}{2} \text{ and } \#W + \#I_Y \geq \frac{1}{2} \tag{9}$$

Inequality (8) is equivalent to the condition $\#L + \min\{\#I_X, \#I_Y\} \geq \frac{1}{2}$, while (9) is equivalent to the condition $\#W + \min\{\#I_X, \#I_Y\} \geq \frac{1}{2}$. The committee has a collective preference for ignorance if and only if at least one of these conditions holds. Hence, the committee has a collective preference for ignorance if and only if

$$\max\{\#W, \#L\} + \min\{\#I_X, \#I_Y\} \geq \frac{1}{2}$$

which is equivalent to $\max\{\#W, \#L\} + \min\{\#I_X, \#I_Y\} \geq \min\{\#W, \#L\} + \max\{\#I_X, \#I_Y\}$. Rearranging, we obtain $\max\{\#I_X, \#I_Y\} - \min\{\#I_X, \#I_Y\} \leq \max\{\#W, \#L\} - \min\{\#W, \#L\}$. This is equivalent to $|\#I_X - \#I_Y| \leq |\#W - \#L|$.

Proof of Proposition 3

If the committee is unable to commit to rejecting the proposal in ignorance, then the voter i's gain from adopting the reform at time 0 instead of waiting equals

$$\hat{d}_i = (\pi_0 x_i + [1-\pi_0] y_i) - \pi_0 x_i g_\gamma(x) - (1-\pi_0) y_i g_\gamma(y)$$

Let $\hat{d} \equiv \left(\hat{d}_i\right)_{i \in I}$ denote the vector of these gains across voters.

1. If $g_\gamma(x) = g_\gamma(y) = 1$, then $\hat{d}_i = 0$, $\forall i \in I$, so all agents are indifferent between learning and not learning.
2. If $g_\gamma(x) = g_\gamma(y) = 0$, then $\hat{d} = \pi_0 x + [1-\pi_0] y$, so the committee has a strict preference for ignorance whenever $g_\gamma(\pi_0 x + [1-\pi_0] y) = 1$, and a strict preference for learning whenever $g_\gamma(\pi_0 x + [1-\pi_0] y) = 0$.
3. If $g_\gamma(x) = 0$ and $g_\gamma(y) = 1$, then $\hat{d} = \pi_0 x$, so $g_\gamma\left(\hat{d}\right) = g_\gamma(x) = 0$, hence the committee chooses to learn
4. If $g_\gamma(x) = 1$ and $g_\gamma(y) = 0$, then $\hat{d} = (1-\pi_0) y$, so $g_\gamma\left(\hat{d}\right) = g_\gamma(y) = 0$, hence the committee chooses to learn.

Proof of Proposition 4

Take the case when $g_\gamma(x) = 1$ and $g_\gamma(y) = 0$. Take any strategy profile, and look at the associated $T(\pi)$. If $\omega = X$, the reform will always be adopted; and if $\omega = Y$, the reform will be adopted if and only if no signal arrives during the time interval $T(\pi)$. Thus, agent i's expected payoff under that strategy profile is

$$\pi x_i + (1-\pi)\,\mathrm{e}^{-\lambda_Y T(\pi)} y_i$$

For any $T(\pi) > 0$, this is smaller that $s_i(\pi)$ if and only if $y_i > 0$. But $g(y) = 0$, and this implies that the number of voters for whom $y_i > 0$ is smaller than γ. Hence, at any $\pi \in (0,1)$, the number of voters in favour of adopting the proposal is less than γ, so the group never votes to stop and continues acquiring information until the state is revealed. It is straightforward to show that the same outcome emerges if $g_\gamma(x) = 0$ and $g_\gamma(y) = 1$.

Proof of Proposition 5

Consider the problem faced by $i \in I_X$. Take any strategy profile under which $T(\pi) > 0$. If $s_i(\pi) > 0$, then voter i's payoff from following this strategy profile and delaying the decision is

$$\pi \mathrm{e}^{-\lambda_X T(\pi)} x_i + (1-\pi)\,\mathrm{e}^{-\lambda_Y T(\pi)} y_i < \pi \mathrm{e}^{-\lambda_Y T(\pi)} x_i + (1-\pi)\,\mathrm{e}^{-\lambda_Y T(\pi)} y_i <$$
$$< \pi x_i + (1-\pi)\, y_i = s_i(\pi_t)$$

and if $s_i(\pi) < 0$, then i's payoff from waiting equals

$$\pi \mathrm{e}^{-\lambda_X T(\pi)} x_i + (1-\pi)\,\mathrm{e}^{-\lambda_Y T(\pi)} y_i > \pi \mathrm{e}^{-\lambda_X T(\pi)} x_i + (1-\pi)\,\mathrm{e}^{-\lambda_X T(\pi)} y_i >$$
$$> \pi x_i + (1-\pi)\, y_i = s_i(\pi_t)$$

Thus, $i \in I_X$ prefers adopting the reform instead of delaying the decision if and only if $s_i(\pi) \geq 0$, i.e. if and only if $\pi \geq \pi_i^M$.

Proof of Proposition 6

Take an $i \in I_Y$. Suppose that as long as the state is not revealed, i can choose when to adopt the reform. Denote by$V_i(\pi, T)$ i's expected payoff from delaying the adoption of reform by some time T, when the current belief is π. Then $V_i(\pi, T) = \pi \mathrm{e}^{-\lambda_X T} x_i +$

$(1-\pi)\,\mathrm{e}^{-\lambda_Y T} y_i$. By continuity, when T is close to zero, $V_i(\pi,T)$ must be increasing in T at all beliefs π at which i prefers to delay the reform; and decreasing at all beliefs π at which i prefers to adopt it. Differentiating, we obtain:

$$\left.\frac{dV_i(\pi,T)}{dT}\right|_{T=0} = -\lambda_X \pi x_i - \lambda_Y (1-\pi)\, y_i$$

which is positive for all $\pi > \frac{\lambda_Y y_i}{\lambda_Y y_i - \lambda_X x_i} = \pi_i^*$, and negative for all $\pi < \pi_i^*$. ∎

Proof of Proposition 7 (Second Case)

Suppose that $g(x) = g(y) = 0$, and take a $\pi \leq \frac{1}{2}$. Consider the equilibrium waiting time $T(\pi)$. If $T(\pi) = \infty$, the proposal is never adopted and the value of all agents is zero irrespective of λ_X, λ_Y. For a finite $T(\pi)$, let $\pi^* \leq \pi$ be the public belief after $T(\pi)$ units of time, if no signal has arrived by that time. By Bayes law, π^* is given by:

$$\pi^* = \frac{\pi \mathrm{e}^{-\lambda_X T}}{\pi \mathrm{e} + (1-\pi)\,\mathrm{e}^{-\lambda_Y T}}$$

which implies that

$$\mathrm{e}^T = \left(\frac{\pi}{1-\pi}\right)^{\frac{1}{\lambda_X - \lambda_Y}} \left(\frac{1-\pi^*}{\pi^*}\right)^{\frac{1}{\lambda_X - \lambda_Y}} = \left(\frac{\pi}{1-\pi}\right)^{\frac{1}{\lambda_X - \lambda_Y}} D$$

where $D \equiv \left(\frac{1-\pi^*}{\pi^*}\right)^{\frac{1}{\lambda_X - \lambda_Y}} \leq 1$, the last inequality coming from the fact that $\pi^* \leq \pi \leq \frac{1}{2}$. Then the value of agent $i = 0$ at a belief π equals:

$$\begin{aligned} u(a,b,\pi) &= \pi \mathrm{e}^{-\lambda_X T} a + (1-\pi)\,\mathrm{e}^{-\lambda_Y T} b = \\ &= \pi \left(\tfrac{\pi}{1-\pi}\right)^{\frac{-\lambda_X}{\lambda_X - \lambda_Y}} D^{-\lambda_X} a + (1-\pi) \left(\tfrac{\pi}{1-\pi}\right)^{\frac{-\lambda_Y}{\lambda_X - \lambda_Y}} D^{-\lambda_Y} b \end{aligned}$$

which can be written as

$$u(a,b,\pi) = C\left(D^{-\lambda_X} a + D^{-\lambda_Y} b\right)$$

where $C \equiv \pi^{\frac{-\lambda_Y}{\lambda_X - \lambda_Y}} (1-\pi)^{\frac{\lambda_X}{\lambda_X - \lambda_Y}} > 0$.

Similarly, the value of agent 1 equals:

$$u(b,a,\pi) = C\left(D^{-\lambda_X} b + D^{-\lambda_Y} a\right)$$

Subtracting, we obtain:

$$u(a,b,\pi) - u(b,a,\pi) = C(a-b)\left(D^{-\lambda_X} - D^{-\lambda_Y}\right) \leq 0$$

where the last inequality is due to the fact that $C > 0$, $a < b$, and $D^{-\lambda_X} \leq D^{-\lambda_Y}$. Thus, the value of agent 0 is greater than that of agent 1.

Hence, an agent's value either does not depend on λ_X, λ_Y , or if it does, then for two agents with symmetric preferences, the agent whose least preferred state is revealed faster has a higher value as long as the probability of that state is less than $\frac{1}{2}$. ∎

Proof of Proposition 9.

Suppose the planner has selected some $\gamma \geq \frac{1}{2}$. Then, when Nature draws s from pdf F, and the committee makes a vote knowing the it, we can have the following situations:

1. $s < 1-\gamma$. This happens with probability $F(1-\gamma)$. In this case, $g(x) = 0$ and $g(y) = 1$, hence $g(v) = 1$. Thus, if $\omega = X$, every agent gets zero, and if $\omega = Y$, agent i recieves -1 if $i \in I_X$ and k if $i \in I_Y$. The expected sum of payoffs, conditional on $s < 1-\gamma$, will equal $\frac{1}{2}\mathrm{E}_F\left[-s + a(1-s) \mid s < 1-\gamma\right]$.

2. $s > \gamma$. This happens with probability $1-F(\gamma)$. In this case, $g(x) = 1$ and $g(y) = 0$, hence $g(v) = 1$. Thus, if $\omega = Y$, every agent gets zero, and if $\omega = X$, agent i recieves a if $i \in I_X$ and -1 if $i \in I_Y$. The expected sum of payoffs, conditional on $s > \gamma$, will equal $\frac{1}{2}\mathrm{E}_F\left[as - (1-s) \mid s > \gamma\right]$.

3. $s \in [1-\gamma, \gamma]$. This happens with probability $F(\gamma) - F(1-\gamma)$. In this case, $g(x) = g(y) = 0$, so hence $g(v) = 0$. When they make the decision in ignorance, every agent's payoff is $\frac{a-1}{2}$ if $a > 1$ and 0 if $a < 1$.

Denote the expected sum of payoffs by $U(\gamma)$. When $a > 1$, it equals:

$$\begin{aligned}
U(\gamma) &= F(1-\gamma)\,\mathrm{E}_F\left[\tfrac{-(a+1)s+a}{2} \mid s < 1-\gamma\right] + \left[1-F(\gamma)\right]\mathrm{E}_F\left[\tfrac{(a+1)s-1}{2} \mid s > \gamma\right] + \\
&+ \left[F(\gamma) - F(1-\gamma)\right]\tfrac{a-1}{2} = \tfrac{1}{2}\left(-(a+1)\int_0^{1-\gamma} s\,dF(s) + aF(1-\gamma)\right) + \\
&+ \tfrac{1}{2}\left((a+1)\int_\gamma^1 s\,dF(s) - \left[1-F(\gamma)\right] + \left[F(\gamma) - F(1-\gamma)\right](a-1)\right)
\end{aligned}$$

and when $a < 1$, it equals

$$U(\gamma) = F(1-\gamma)\,\mathrm{E}_F\left[\tfrac{-(a+1)s+a}{2} \mid s < 1-\gamma\right] + [1-F(\gamma)]\,\mathrm{E}_F\left[\tfrac{(a+1)s-1}{2} \mid s > \gamma\right] =$$
$$= \tfrac{1}{2}\Big(-(a+1)\int_0^{1-\gamma} sdF(s) + aF(1-\gamma) + (a+1)\int_\gamma^1 sdF(s) - [1-F(\gamma)]\Big)$$

If $a > 1$, $\frac{dU(\gamma)}{d\gamma} = \frac{1}{2}[1-\gamma-a\gamma][f(1-\gamma)+f(\gamma)]$. This expression is positive for $\gamma < \frac{1}{1+a}$ and negative for $\gamma > \frac{1}{1+a}$. Hence, $\gamma^* = \frac{1}{1+a}$.

If $a < 1$, $\frac{dU(\gamma)}{d\gamma} = \frac{1}{2}[a-\gamma-a\gamma][f(1-\gamma)+f(\gamma)]$. This expression is positive for $\gamma < \frac{a}{1+a}$ and negative for $\gamma > \frac{a}{1+a}$. Hence, $\gamma^* = \frac{a}{1+a}$.■

Proof of Proposition 10

If the planner has chosen $\gamma \geq \frac{1}{2}$, then from her point of view, the expected payoffs are as follows:

1. If $s < 1-\gamma$ - this happens with probability $F(1-\gamma)$ - the committee learns the state and adopts the reform iff $\omega = Y$. The ex ante expected sum of payoffs in that case is $\frac{1}{2}\mathrm{E}[-s+a(1-s) \mid s < 1-\gamma]$.
2. If $s > \gamma$ - this happens with probability $[1-F(\gamma)]$ - the committee learns the state and adopts the reform iff $\omega = X$. The ex ante expected sum of payoffs is $\frac{1}{2}\mathrm{E}[as-(1-s) \mid s > \gamma]$.
3. If $1-\gamma < s < \gamma$ - this happens with probability $F(\gamma) - F(1-\gamma)$ - the committee chooses not to learn the state, and, as every voter's expected payoff from the reform is $\frac{a+1}{2}$, the reform is adopted in either state iff $a > 1$. The expected sum of payoffs is $\frac{1}{2}\mathrm{E}[a-1 \mid 1-\gamma < s < \gamma,\, a > 1]$.

Then the expected utilitarian welfare equals:

$$U(\gamma) = F(1-\gamma)\,\tfrac{1}{2}\mathrm{E}[-s+a(1-s) \mid s < 1-\gamma] + [1-F(\gamma)]\,\tfrac{1}{2}\mathrm{E}[as-(1-s) \mid s > \gamma] +$$
$$+ [F(\gamma)-F(1-\gamma)]\,\tfrac{1}{2}\Pr(a>1)(\mathrm{E}[a \mid a>1]-1) =$$
$$= \tfrac{1}{2}\Big(-(\mathrm{E}[a]+1)\int_0^{1-\gamma} sdF(s) + \mathrm{E}_H[a]\,F(1-\gamma)\Big) +$$
$$+ \tfrac{1}{2}\Big((\mathrm{E}_H[a]+1)\int_\gamma^1 sdF(s) - [1-F(\gamma)] + \Pr(a>1)[F(\gamma)-F(1-\gamma)](\mathrm{E}_H[a \mid a>1]-1)\Big)$$

where $p_1 \equiv \Pr(a>1)$, and $k_1 \equiv \mathrm{E}[a \mid a>1]$. Then

$$\frac{dU}{d\gamma} = \frac{1}{2}[(1-\gamma-\mathrm{E}_H[a]\,\gamma) + \Pr(a>1)(\mathrm{E}_H[a \mid a>1]-1)][f(1-\gamma)+f(\gamma)]$$

which is positive if $\gamma < \frac{1+\Pr(a>1)(\mathrm{E}_H[a|a>1]-1)}{(\mathrm{E}_H[a]+1)}$, and negative if $\gamma > \frac{1+\Pr(a>1)(\mathrm{E}_H[a|a>1]-1)}{(\mathrm{E}_H[a]+1)}$. Hence $\gamma^* = \min\left\{\frac{1+\Pr(a>1)(\mathrm{E}_H[a|a>1]-1)}{(\mathrm{E}_H[a]+1)}, 1\right\}$.

Proof of Proposition 11

If $g_{0.5}(x) = g_{0.5}(y) = g_{0.5}(\pi_0 x + [1-\pi_0] y)$, then the decision on the reform is the same wiht or without information, so commitment to learning has no effect. If $g_{0.5}(x) \neq g_{0.5}(y)$, the committee chooses to learn the state anyway, so a commitment to learning again has no effect. The only case when it does have an effect is when $g_{0.5}(x) = g_{0.5}(y) \neq g_{0.5}(\pi_0 x + [1-\pi_0] y)$.

Suppose that $g_{0.5}(x) = g_{0.5}(y) = 1$ and $g_{0.5}(\pi_0 x + [1-\pi_0] y) = 0$. Then, $d(\pi_0 x + [1-\pi_0] y) < 0$. Without a commitment to learning, the committee votes in ignorance and rejects the reform, giving each member a payoff of zero. With a commitment to learning, the reform is adopted in either state, so the expected payoff of each voter i is $\pi_0 x_i + [1-\pi_0] y_i$. Commitment to learning is then socially optimal iff $w(\pi_0 x + [1-\pi_0] y) > 0$.

Now suppose that $g_{0.5}(x) = g_{0.5}(y) = 0$ and $g_{0.5}(\pi_0 x + [1-\pi_0] y) = 1$, hence $d(\pi_0 x + [1-\pi_0] y) > 0$. Without a commitment to learning, the committee votes in ignorance and adopts the reform, giving each voter i an expected payoff of $\pi_0 x_i + [1-\pi_0] y_i$. With a commitment to learning, the reform is rejected in either state, and the payoff of each voter is 0. Commitment to learning is then socially optimal iff $0 > w(\pi_0 x + [1-\pi_0] y)$.

Hence, whenever $sign[d(\pi_0 x + [1-\pi_0] y)] \neq sign[w(\pi_0 x + [1-\pi_0] y)]$, commitment to learning either has no effect, or is socially preferable. Similarly, when $sign[d(\pi_0 x + [1-\pi_0] y)] =\neq sign[w(\pi_0 x + [1-\pi_0] y)]$,commitment to learning either has no effect, or is socially harmufl.

Proof of Proposition 12

The value of ignorance for agent i is

$$\begin{gathered}
d_i = (\pi_0 x_i + [1-\pi_0] y_i)\, g_\gamma(\pi_0 x + [1-\pi_0] y) - \\
-(\pi_0 p x_i + [1-\pi_0] q y_i)\, g_\gamma(\pi_0 p x + [1-\pi_0] q y) - \\
-(\pi_0 [1-p] x_i + [1-\pi_0][1-q] y_i)\, g_\gamma(\pi_0 [1-p] x + [1-\pi_0][1-q] y)
\end{gathered}$$

1. If $g_\gamma(\pi_0 p x + [1-\pi_0] q y) = g_\gamma(\pi_0 [1-p] x + [1-\pi_0][1-q] y) = g_\gamma(\pi_0 x + [1-\pi_0] y)$, then $d_i = 0$, $\forall i \in I$, so all agents are indifferent between learning and not learning.

2. If $g_\gamma(\pi_0 px + [1-\pi_0]qy) = g_\gamma(\pi_0[1-p]x + [1-\pi_0][1-q]y) = 0$ and $g_\gamma(\pi_0 x + [1-\pi_0]y) = 1$, then $d = \pi_0 x + [1-\pi_0]y$, so $g_\gamma(d) = g(\pi_0 x + [1-\pi_0]y) = 1$.

3. If $g_\gamma(\pi_0 px + [1-\pi_0]qy) = g_\gamma(\pi_0[1-p]x + [1-\pi_0][1-q]y) = 1$ and $g_\gamma(\pi_0 x + [1-\pi_0]y) = 0$, then $d = -(\pi_0 x + [1-\pi_0]y)$, so $g_\gamma(d) = 1 - g_\gamma(\pi_0 x + [1-\pi_0]y) = 1$.

4. If $g_\gamma(\pi_0 px + [1-\pi_0]qy) = 1$ and $g_\gamma(\pi_0[1-p]x + [1-\pi_0][1-q]y) = g_\gamma(\pi_0 x + [1-\pi_0]y) = 0$, then $d = -(\pi_0 px + [1-\pi_0]qy)$, so $g_\gamma(d) = 1 - g_\gamma(-d) = 0$.

5. In a similar way, it can be shown that $g_\gamma(\pi_0[1-p]x + [1-\pi_0][1-q]y) = 1$ and $g_\gamma(\pi_0 px + [1-\pi_0]qy) = g_\gamma(\pi_0 x + [1-\pi_0]y) = 0$ imply $g_\gamma(d) = 0$.

6. If $g_\gamma(\pi_0 px + [1-\pi_0]qy) = 0$ and $g_\gamma(\pi_0[1-p]x + [1-\pi_0][1-q]y) = g_\gamma(\pi_0 x + [1-\pi_0]y) = 1$, then $d = \pi_0 px + [1-\pi_0]qy$, so $g_\gamma(d) = g_\gamma(\pi_0 px + [1-\pi_0]qy) = 0$.

7. In a similar way, it can be shown that $g_\gamma(\pi_0[1-p]x + [1-\pi_0][1-q]y) = 0$ and $g_\gamma(\pi_0 px + [1-\pi_0]qy) = g_\gamma(\pi_0 x + [1-\pi_0]y) = 1$ imply $g_\gamma(d) = 0$.∎

Proof of Proposition 13

The value of ignorance to agent i - the difference in her utilities from acquiring information structure $\mathcal{P}$ and from remaining ignorant - equals:

$$d_i = \sum_{j\in\Omega} p_i^j x_i^j g_{0.5}\left[\sum_{j\in\Omega} p^j x^j\right] - \sum_{S\in\mathcal{P}}\left(\sum_{j\in S} p_i^j x_i^j g_{0.5}\left[\sum_{j\in S} p^j x^j\right]\right)$$

Or:

$$d_i = \sum_{S\in P}\sum_{j\in S}\left[p_i^j x_i^j\left(g_{0.5}\left[\sum_{j\in\Omega} p^j x^j\right] - g_{0.5}\left[\sum_{j\in S} p^j x^j\right]\right)\right]$$

By the definition of M, the expression in the brackets equals 0 for all $S \notin M$. For all $S \in M$, the expression in the brackets equals $-p_i^j x_i^j$ if $g_{0.5}\left[\sum_{j\in\Omega} p^j x^j\right] = 0$, and it equals $p_i^j x_i^j$ if $g_{0.5}\left[\sum_{j\in\Omega} p^j x^j\right] = 1$. Hence, if $g_{0.5}\left[\sum_{j\in\Omega} p^j x^j\right] = 0$, then $d_i = \sum_{j\in S\in M}\left(-p_i^j x_i^j\right)$, and hence $g_{0.5}(d) = 0$ iff $g_{0.5}\left[\sum_{j\in S\in M} p^j x^j\right] = 1$. Similarly, if $g_{0.5}\left[\sum_{j\in\Omega} p^j x^j\right] = 1$, then $d_i = \sum_{j\in S\in M}\left(p_i^j x_i^j\right)$, and hence $g_{0.5}(d) = 0$ iff $g_{0.5}\left[\sum_{j\in S\in M} p^j x^j\right] = 0$.∎

Proof of Proposition 15

By Bayes' law, we have

$$\Pr\left(y=0 \mid s_i=0\,,\,piv\right)=\frac{\frac{1}{2}p\begin{pmatrix}k\\n\end{pmatrix}q_0^k\left(1-q_0\right)^{n-k}}{\frac{1}{2}p\begin{pmatrix}k\\n\end{pmatrix}q_0^k\left(1-q_0\right)^{n-k}+\frac{1}{2}\left(1-p\right)\begin{pmatrix}k\\n\end{pmatrix}q_1^k\left(1-q_1\right)^{n-k}}$$

and

$$\Pr\left(y=0 \mid s_i=1\,,\,piv\right)=\frac{\frac{1}{2}\left(1-p\right)\begin{pmatrix}k\\n\end{pmatrix}q_0^k\left(1-q_0\right)^{n-k}}{\frac{1}{2}\left(1-p\right)\begin{pmatrix}k\\n\end{pmatrix}q_0^k\left(1-q_0\right)^{n-k}+\frac{1}{2}p\begin{pmatrix}k\\n\end{pmatrix}q_1^k\left(1-q_1\right)^{n-k}}$$

Then we have:

$$\mathrm{E}\left[u\left(0,y\right)\mid 0\,,\,piv\right]=\frac{pa-\left(1-p\right)M}{p+\left(1-p\right)M} \tag{10}$$

$$\mathrm{E}\left[u\left(1,y\right)\mid 0\,,\,piv\right]=\frac{-p+\left(1-p\right)Ma}{p+\left(1-p\right)M} \tag{11}$$

$$\mathrm{E}\left[u\left(0,y\right)\mid 1\,,\,piv\right]=\frac{\left(1-p\right)a-pM}{\left(1-p\right)+pM} \tag{12}$$

$$\mathrm{E}\left[u\left(1,y\right)\mid 1\,,\,piv\right]=\frac{-\left(1-p\right)+pMa}{\left(1-p\right)+pM} \tag{13}$$

The conditions under which (10)-(13) are greater than zero are equivalent to the conditions in Proposition 15.

Proof of Proposition 16.

To show that Proposition 16 holds, note that (3.1) is continuous in all parameters, and consider its limit properties.

When $\gamma \to \frac{1}{2}$, (3.1) becomes

$$\max\left\{\frac{1-p}{p}\frac{1}{a}\,,\,\frac{1-p}{p}a\right\}\leq 1\leq \min\left\{\frac{p}{1-p}\frac{1}{a}\,,\,\frac{p}{1-p}a\right\}$$

When $k \to \frac{n}{2}$, (3.1) becomes

$$\max\left\{\frac{1-p}{p}\frac{1}{a}, \frac{1-p}{p}a\right\} \leq 1 \leq \min\left\{\frac{p}{1-p}\frac{1}{a}, \frac{p}{1-p}a\right\}$$

When $p \to 1$, (3.1) becomes

$$0 \leq \left(\frac{\gamma}{1-\gamma}\right)^{2k-n} \leq \infty$$

And each of these holds for all values of the remaining parameters, assuming that $a \in \left(\frac{1-p}{p}, \frac{p}{1-p}\right)$.

Proof of Proposition 17

To prove (1), take a set S. Now take a state w belonging to the interior of S and suppose that $z_S(w) < 0$.[5] We want to prove that S is not optimal.

Consider a deviation from S to $\hat{S} \equiv S \setminus [w, r]$. If $r = w$, then $v\left(\hat{S}\right) = v(S)$. Recall that $v(S) = \int\limits_{\omega \notin S} F(\omega)\, g(\omega)\, d\omega + \mu_S F(t_S)$. Then

$$v\left(\hat{S}\right) - v(S) = F\left(t_{S\setminus[w,r]}\right)\mu_{S\setminus[w,r]} + \int_w^r F(\omega)g(\omega)d\omega - F(t_S)\,\mu_S$$

Note that that

$$\mu_{S\setminus[w,r]} = \int\limits_{\omega\in S} g(\omega)\, d\omega - \int_w^r g(\omega)\, d\omega$$

and

$$t_{S\setminus[w,r]} = \frac{\int\limits_{\omega\in S} \omega g(\omega)\, d\omega - \int_w^r \omega g(\omega)\, d\omega}{\int\limits_{\omega\in S} g(\omega)\, d\omega - \int_w^r g(\omega)\, d\omega}$$

Taking the derivative of $v\left(\hat{S}\right) - v(S)$ with respect to r yields

$$\frac{\partial\left[v\left(\hat{S}\right) - v(S)\right]}{\partial r} = g(r)\left[f\left(t_{S\setminus[w,r]}\right)\left(t_{S\setminus[w,r]} - r\right) - F\left(t_{S\setminus[w,r]}\right) + F(r)\right] = -g(r)\, z_{S\setminus[w,r]}(r)$$

[5] The assumption that w is in the interior of A is without loss of generality, since for every w on the boundary of A such that $z_S(w) < 0$, there must - since $z_A(w)$ is continuous - be another w' in the neighborhood of w for which $z_S(w)$ is also negative.

If $r = w$, then $S \setminus [w, r] = S$, and $\left[v\left(\hat{S}\right) - v\left(S\right)\right] = 0$. If S is an optimal strategy, that difference must be weakly decreasing in r at $r = w$. But if $z_S\left(w\right) < 0$, then

$$\left.\frac{\partial\left[v\left(\hat{S}\right) - v\left(S\right)\right]}{\partial r}\right|_{r=w} = -g\left(w\right) z_S\left(w\right) > 0$$

as g is assumed to be strictly positive everywhere. Therefore, the Sender benefits from increasing r, which means that S is not optimal.

Part (2) is proved analogously. Suppose that $z_S\left(w\right) > 0$ for some $w \notin S$. Now take some interval $[w, r]$ such that every $\omega \in [w, r]$ is a singleton element of the partition, and consider a change from S to $\hat{S} \equiv S \bigcup [w, r]$. Then

$$\left.\frac{\partial\left[v\left(\hat{S}\right) - v\left(S\right)\right]}{\partial r}\right|_{r=w} = g\left(w\right) z_A\left(w\right) > 0$$

so again there is a profitable deviation.

Proof of Corollary 4.3.1

To prove necessity, suppose that F is not convex - this implies that it is strictly concave on some interval $[p, q]$. Consider a deviation from a full revelation strategy to a strategy in which $S = [p, q]$. Then the change between the government's payoffs from such deviation equals:

$$\begin{aligned}
&v\left(\emptyset\right) - v\left([p, q]\right) = F\left(t_{[p,q]}\right)\mu_{[p,q]} - \textstyle\int_p^q F\left(\omega\right) g\left(\omega\right) d\omega = \\
&= F\left[\mathrm{E}\left(\omega \mid \omega \in [p, q]\right)\right] \Pr\left(\omega \in [p, q]\right) - \mathrm{E}\left(F\left[\omega\right] \mid \omega \in [p, q]\right) \Pr\left(\omega \in [p, q]\right) = \\
&= \Pr\left(\omega \in [p, q]\right)\left[F\left[\mathrm{E}\left(\omega \mid \omega \in [p, q]\right)\right] - \mathrm{E}\left(F\left[\omega\right] \mid \omega \in [p, q]\right)\right] > 0
\end{aligned}$$

where the last inequality sign follows from Jensen's inequality. Hence, if F is not weakly convex, full disclosure cannot be optimal.

Proof of Proposition 18

Consider a censorship policy $S = \bigcup_{i=1}^{n} [p_j, q_j]$ such that $0 \leq a_j < b_j < a_{j+1} \leq 1$, $\forall j \in \{1, ...n\}$. What is the smallest number of local maxima that f needs to have for S to be the Sender's equilibrium strategy?

Observe that since we have assumed the number of weak local maxima to be finite, f cannot be constant at any interval. This means that $z_S(\cdot)$ cannot equal zero on any interval $[r,s] \subseteq [0,1]$, since if it was zero, this would mean that $f(\omega) = f(t_S)$, $\forall \omega \in [r,s]$, i.e. that f is horizontal. The fact that $z_S(\cdot)$ cannot be zero on an interval, together with Proposition 17, implies that $z_S(\omega)$ is increasing at $\omega = p_j$ and decreasing at $\omega = q_j$, $\forall j \in \{1,...n\}$. This means - since $z_S(\omega)$ is continuously differentiable - that $\frac{dz_S}{d\omega}(p_j) > 0$ and $\frac{dz_S}{d\omega}(q_j) < 0$.

Hence, for every $j \in \{1,...n\}$, $z_S(\omega)$ must have at least one local maximum $c_j \in (p_j, q_j)$ and at least one local minimum $d_j \in (q_j, p_{j+1})$. At a local maximum, $\frac{dz_S^2}{d\omega^2}(c_j) = -f'(c_j) < 0$, while at a local minimum, $\frac{dz_S^2}{d\omega^2}(d_j) = -f'(d_j) > 0$. But f is assumed to be continuously differentiable, and thus for every $j \in \{1,...n\}$ there must be a state $w_j \in (c_j, d_j)$ such that f' is positive to the left of w_j and negative to the right of it. This w_j is therefore a local maximum of f, and there must be such a point in every interval (p_j, p_{j+1}). There are $n-1$ such intervals, which gives us $n-1$ local maxima.

To see that another maximum of f must exist, note that $z_S(t_S) = 0$, and also, $\frac{dz_S}{d\omega}(t_S) = 0$. This gives several possibilities. If $t_S \in (q_j, p_{j+1})$ for some $j \in \{1,...n\}$, then it is a local maximum of z_S, and there are not one but at least two local minima in (q_j, p_{j+1}) - one to the left of t_S and one to the right. So there is one more pair of a maximum and a minimum of $z_S(\omega)$, and by the above logic, there must be a local maximum of f in addition to the ones found in the previous paragraph. If $t_S \in (p_j, q_j)$ for some $j \in \{1,...n\}$, then it is a local minimum, and there are two local maxima in (p_j, q_j) - again, f must have an extra local maximum. If $t_S = p_j$ for some $j \in \{1,...n\}$, then the shape of $z_S(\omega)$ implies that in some neighbourhood of p_j, $\frac{dz_S^2}{d\omega^2}(\omega) = -f'(\omega) < 0$ for $\omega < p_j$ and $\frac{dz_S^2}{d\omega^2}(\omega) = -f'(\omega) > 0$ for $\omega > p_j$. Consequently, f must have an additional local maximum at p_j. Finally, if $t_S = q_j$ for some $j \in \{1,...n\}$, then the shape of $z_S(\omega)$ implies that f must have not one but at least two local maxima in (c_j, d_j).

Hence, if S forms part of an equilibrium, f must have at least $n-1$ local maxima plus one more. Therefore, $m \geq n$.∎

Proof of Proposition 19

Part 1

From Proposition 18 and Corollary 4.3.1 it follows that under a unimodal density f with a peak on $(0,1)$, the set S will consist of exactly one interval. Therefore, t_S must be in the interior of S. If $t_S \leq k$, this would mean that f is increasing on some neighbourhood of t_S, implying that $z_S(\omega) = \int_{t_S}^{\omega} f(t_S)\,dx - \int_{t_S}^{\omega} f(x)\,dx < 0$ for states in that neighborhood. Since this neighbourhood belongs to S, by Proposition 2 this cannot hold at an equilibrium. Thus, $t_S \in (k,1]$. Then $z_S(\omega) > 0$ for all $\omega > t_S$. Similarly, $z_S(k) > 0$, and $z_S(\omega) > 0$ for some $\omega < k$. Hence, $S = [p,1]$ for some $p \in [0,k)$.

Part 2

Take any set $S \subseteq [0,1]$. If $t_S \in (k,1)$, then $z_S(\omega) < 0$ for all $\omega > t_S$. Hence, at an equilibrium it must be that all $\omega > t_S$ do not belong to S. But this cannot be hold, since $t_S \equiv \mathrm{E}[\omega \mid \omega \in S]$. Hence, at an equilibrium, S is such that $t_S \in [0,k]$. Then $z_S(\omega)$ is positive for all $\omega < k$, and is decreasing on $[k,1]$. Thus, $S = [0,q]$ for some $q \in (k,1]$.

Proof of Proposition 22.

Consider symmetric strategies $h_1 = h_2 = ... = h_n \equiv h$. Denote by $z(h)$ the probability that a contestant wins the prize if he does not take the test; note that this probability does not depend the contestant's type. Denote by $\pi(x,h)$ be the probability of winning the prize for a contestant with type x who takes the test. Then a contestant with type x weakly prefers to take the test whenever $\pi(x,h) - c \geq z(h)$, and weakly prefers not taking it whenever $\pi(x,h) - c \leq z(h)$.

Denote by b_n the highest type for which the probability of taking the test is below 1. Take a type $\hat{x} < b_n$. We must have $\pi(b_n,h) - c \leq z(h)$, but also $\pi(\hat{x},h) - c \geq z(h)$. But $\pi(x,h)$ is nondecreasing in x, so $\pi(b_n,h) = \pi(\hat{x},h)$.

Note that $\pi(b_n,h) = \left(1 - \int_b^1 h(u)\,dF(u)\right)^{n-1}$. Let m be the expected value of a contestant's type, conditional on him not taking the test. If $\hat{x} < m$, then $\pi(\hat{x},h) = \left(1 - \int_{\hat{x}}^1 h(u)\,dF(u) - \int_0^1 [1-h(u)]\,dF(u)\right)^{n-1}$. If $\hat{x} \geq m$, then $\pi(\hat{x},h) = \left(1 - \int_{\hat{x}}^1 h(u)\,dF(u)\right)^{n-1}$. Either way, for $\pi(b_n,h) = \pi(\hat{x},h)$ to hold, it must be that $\int_{\hat{x}}^{b_n} h(u)\,dF(u) = 0$. This should be true for any $\hat{x} < b_n$. Thus, $\Pr[h(x) > 0 \mid x_i < b_n] = 0$.

Proof of Proposition 28.

If $F(x) = x$, then the gain from making the test mandatory is given by

$$\Delta = b_n^{n-1} \int_0^{b_n} x dx - \int_{\hat{b}_n}^{b_n} x^n dx - \hat{b}_n^{n+1} = \frac{1}{2} b_n^{n+1} - \frac{1}{n+1} b_n^{n+1} + \frac{1}{n+1} \hat{b}_n^{n+1} - \hat{b}_n^{n+1}$$

or

$$\Delta = \frac{n-1}{2(n+1)} b_n^{n+1} - \frac{n}{n+1} \hat{b}_n^{n+1}$$

which after plugging in the expressions for b_n and $\hat{b}_n$ and simplifying equals

$$\Delta = \frac{1}{n+1} c^{\frac{n+1}{n-1}} \left[\frac{n-1}{2} \left(\frac{n}{n-1} \right)^{\frac{n+1}{n-1}} - n \right]$$

which has the same sign as $\frac{n-1}{2} \left(\frac{n}{n-1} \right)^{\frac{n+1}{n-1}} - n$, or as $\frac{1}{2} \left(\frac{n}{n-1} \right)^{\frac{2}{n-1}} - 1$. But note that for all $n \geq 3$, we have $\left(\frac{n}{n-1} \right)^{\frac{2}{n-1}} < 2$, and thus $\Delta < 0$.

Proof of Proposition 29.

If $b_n = 1$, changing σ does not, in general, affect b_n. Now let us look at what happens if $b_n < 1$.

Recall that $\gamma \equiv \Pr[x_j > b_n \cap x_j + \varepsilon_j < b_n + \varepsilon_i] = \Pr[x_j > b_n \cap \varepsilon_j - \varepsilon_i < -(x_j - b_n)]$. Note that $\varepsilon_j - \varepsilon_i \sim N(0, 2\sigma^2)$. Thus $\gamma = \int_{b_n}^1 f(x) \Phi\left(-\frac{x-b_n}{\sqrt{2}\sigma}\right) dx$. Differentiating with respect to σ yields:

$$\frac{d\gamma}{d\sigma} = -f(b_n)\Phi(0)\frac{db_n}{d\sigma} + \int_{b_n}^1 f(x)\phi\left(-\frac{x-b_n}{\sqrt{2}\sigma}\right)\left(-\frac{-\sqrt{2}\sigma\frac{db_n}{d\sigma} - \sqrt{2}(x-b_n)}{2\sigma^2}\right)dx$$

or

$$\frac{d\gamma}{d\sigma} = -\frac{1}{2}f(b_n)\frac{db_n}{d\sigma} + K\frac{db_n}{d\sigma} + L$$

where $K \equiv \frac{1}{\sqrt{2}\sigma}\int_{b_n}^1 f(x)\phi\left(\frac{x-b_n}{\sqrt{2}\sigma}\right)dx$ and $L \equiv \frac{1}{\sqrt{2}\sigma^2}\int_{b_n}^1 f(x)\phi\left(\frac{x-b_n}{\sqrt{2}\sigma}\right)(x-b_n)\,dx$.

To see the effect of changing σ on b_n, we can differentiate (5.6) with respect to σ, using the above expression for $\frac{d\gamma}{d\sigma}$. This yields:

$$(n-1)\left[F(b_n)+\gamma\right]^{n-2}\left[f(b_n)\frac{db_n}{d\sigma} + \frac{d\gamma}{d\sigma}\right] = \frac{n-1}{n}F(b_n)^{n-2}f(b_n)\frac{db_n}{d\sigma}$$

or

$$\left[F(b_n)+\gamma\right]^{n-2}\left[\frac{1}{2}f(b_n)\frac{db_n}{d\sigma} + K\frac{db_n}{d\sigma} + L\right] = \frac{1}{n}F(b_n)^{n-2}f(b_n)\frac{db_n}{d\sigma}$$

and this gives us:

$$\frac{db_n}{d\sigma} = \frac{\left[F(b_n)+\gamma\right]^{n-2}L}{\frac{1}{n}F(b_n)^{n-2}f(b_n) - \frac{1}{2}\left[F(b_n)+\gamma\right]^{n-2}f(b_n) - \left[F(b_n)+\gamma\right]^{n-2}K}$$

Note that $\gamma > 0$ and $\frac{1}{n} \leq \frac{1}{2}$, and hence, $\frac{1}{n}F(b_n)^{n-2} f(b_n) < \frac{1}{2}[F(b_n) + \gamma]^{n-2} f(b_n)$. Furthermore, $L > 0$ and $K > 0$. Thus, $\frac{db_n}{d\sigma} < 0$.■

Bibliography

Acharya, A. and Meirowitz, A. (2014), 'Sincere voting in large elections', *Mimeo* .

Albrecht, J., Anderson, A. and Vroman, S. (2010), 'Search by committee', *Journal of Economic Theory* **145**(4), 1386–1407.

Alesina, A., Devleeschauwer, A., Easterly, W., Kurlat, S. and Wacziarg, R. (2003), 'Fractionalization', *Journal of Economic Growth* **8**(2), 155–194.

Anderson, C. J. and Paskeviciute, A. (2006), 'How ethnic and linguistic heterogeneity influence the prospects for civil society: A comparative study of citizenship behavior', *Journal of Politics* **68**(4), 783–802.

Anderson, S. P. and McLaren, J. (2012), 'Media mergers and media bias with rational consumers', *Journal of the European Economic Association* **10**(4), 831–859.

Austen-Smith, D. and Banks, J. S. (1996), 'Information aggregation, rationality, and the condorcet jury theorem', *American Political Science Review* **90**(01), 34–45.

Baron, D. P. (2006), 'Persistent media bias', *Journal of Public Economics* **90**(1), 1–36.

Bernhardt, D., Krasa, S. and Polborn, M. (2008), 'Political polarization and the electoral effects of media bias', *Journal of Public Economics* **92**(5), 1092–1104.

Besley, T. and Prat, A. (2006), 'Handcuffs for the grabbing hand? media capture and government accountability', *The American Economic Review* pp. 720–736.

Bhattacharya, S. (2013*a*), 'Preference monotonicity and information aggregation in elections', *Econometrica* **81**(3), 1229–1247.
URL: *http://dx.doi.org/10.3982/ECTA8311*

Bhattacharya, S. (2013*b*), 'Preference monotonicity and information aggregation in elections', *Econometrica* **81**(3), 1229–1247.

Bhattacharya, S. and Mukherjee, A. (2013), 'Strategic information revelation when experts compete to influence', *The RAND Journal of Economics* **44**(3), 522–544.

Board, O. (2009), 'Competition and disclosure', *The Journal of Industrial Economics* **57**(1), 197–213.

Boas, T. C. (2005), 'Television and neopopulism in latin america: Media effects in brazil and peru', *Latin American Research Review* **40**(2), 27–49.

Bouton, L. and Castanheira, M. (2012), 'One person, many votes: Divided majority and information aggregation', *Econometrica* **80**(1), 43–87.

Brennan, G. and Hamlin, A. (1998), 'Expressive voting and electoral equilibrium', *Public Choice* **95**(1-2), 149–175.

Bueno de Mesquita, B. and Downs, G. (2005), 'Development and democracy', *Foreign Aff.* **84**, 77.

Cheong, I. and Kim, J.-Y. (2004), 'Costly information disclosure in oligopoly', *The Journal of Industrial Economics* **52**(1), 121–132.

Chiang, C.-F. and Knight, B. (2011), 'Media bias and influence: Evidence from newspaper endorsements', *The Review of Economic Studies* p. rdq037.

Collier, P. (2001), 'Implications of ethnic diversity', *Economic policy* **16**(32), 127–166.

Compte, O. and Jehiel, P. (2010), 'Bargaining and majority rules: A collective search perspective', *Journal of Political Economy* **118**(2), 189–221.

Condorcet, M. d. (1976), 'Essay on the application of mathematics to the theory of decision-making', *Reprinted in Condorcet: Selected Writings, Keith Michael Baker, ed* **33**.

Damiano, E., Hao, L. and Suen, W. (2011), 'Sincerely voting in large elections', *Mimeo* .

Degan, A. and Merlo, A. (2009), 'Do voters vote ideologically?', *Journal of Economic Theory* **144**(5), 1868–1894.

DellaVigna, S. and Kaplan, E. D. (2007), 'The fox news effect: Media bias and voting', *The Quarterly Journal of Economics* **122**(3), 1187–1234.

Dewan, T. and Myatt, D. P. (2008), 'The qualities of leadership: Direction, communication, and obfuscation', *American Political Science Review* **102**(03), 351–368.

Dewatripont, M. and Roland, G. (1995), 'The design of reform packages under uncertainty', *The American Economic Review* pp. 1207–1223.

Duggan, J. and Martinelli, C. (2011), 'A spatial theory of media slant and voter choice', *The Review of Economic Studies* **78**(2), 640–666.

Dyczok, M. (2006), 'Was kuchma's censorship effective? mass media in ukraine before 2004', *Europe-Asia Studies* **58**(2), 215–238.
URL: *http://www.tandfonline.com/doi/abs/10.1080/09668130500481386*

Easterly, W. and Levine, R. (1997), 'Africa's growth tragedy: policies and ethnic divisions', *The Quarterly Journal of Economics* pp. 1203–1250.

Edmond, C. (2013), 'Information manipulation, coordination, and regime change', *The Review of Economic Studies* p. rdt020.

Egorov, G., Guriev, S. and Sonin, K. (2009), 'Why resource-poor dictators allow freer media: a theory and evidence from panel data', *American Political Science Review* **103**(04), 645–668.

Enikolopov, R., Petrova, M. and Zhuravskaya, E. (2011), 'Media and political persuasion: Evidence from russia', *American Economic Review* **101**(7), 3253–85.
URL: *http://ideas.repec.org/a/aea/aecrev/v101y2011i7p3253-85.html*

Feddersen, T. and Pesendorfer, W. (1997), 'Voting behavior and information aggregation in elections with private information', *Econometrica: Journal of the Econometric Society* pp. 1029–1058.

Fernandez, R. and Rodrik, D. (1991), 'Resistance to reform: Status quo bias in the presence of individual-specific uncertainty', *The American Economic Review* pp. 1146–1155.

Forand, J. G. (2013), 'Competing through information provision', *International Journal of Industrial Organization* **31**(5), 438–451.

Gehlbach, S. and Sonin, K. (2014), 'Government control of the media', *Journal of Public Economics* **118**, 163–171.

Gentzkow, M. and Kamenica, E. (2011), Competition in persuasion, Technical report, National Bureau of Economic Research.

Gentzkow, M. and Kamenica, E. (2015), 'Information environments and the impact of competition on information provision'.

Gentzkow, M. and Shapiro, J. M. (2006), 'Media bias and reputation', *Journal of Political Economy* **114**(2).

Gerardi, D. and Yariv, L. (2007), 'Deliberative voting', *Journal of Economic Theory* **134**(1), 317–338.

Gerardi, D. and Yariv, L. (2008), 'Information acquisition in committees', *Games and Economic Behavior* **62**(2), 436–459.

Gerber, A. S., Karlan, D. and Bergan, D. (2009), 'Does the media matter? a field experiment measuring the effect of newspapers on voting behavior and political opinions', *American Economic Journal: Applied Economics* **1**(2), 35–52.

Gersbach, H. (1991), 'The value of public information in majority decisions', *Economics Letters* **35**(3), 239–242.

Gersbach, H. (2000), 'Public information and social choice', *Social Choice and Welfare* **17**(1), 25–31.

Gersbach, H. and Hahn, V. (2012), 'Information acquisition and transparency in committees', *International Journal of Game Theory* **41**(2), 427–453.

Gershkov, A. and Szentes, B. (2009), 'Optimal voting schemes with costly information acquisition', *Journal of Economic Theory* **144**(1), 36–68.

Giovannoni, F. and Seidmann, D. J. (2007), 'Secrecy, two-sided bias and the value of evidence', *Games and Economic Behavior* **59**(2), 296–315.
URL: *http://ideas.repec.org/a/eee/gamebe/v59y2007i2p296-315.html*

Godefroy, R. and Perez-Richet, E. (2013), 'Choosing choices: Agenda selection with uncertain issues', *Econometrica* **81**(1), 221–253.

Goertz, J. M. and Maniquet, F. (2011), 'On the informational efficiency of simple scoring rules', *Journal of Economic Theory* **146**(4), 1464–1480.

Goldhammer, A. and Elster, J. (2011), *Tocqueville: The Ancien Rme and the French Revolution*, Cambridge University Press.

Goldstein, R. J. (2000), *The war for the public mind: political censorship in nineteenth-century Europe*, Praeger Publishers.

Grossman, S. J. (1981), 'The informational role of warranties and private disclosure about product quality', *Journal of Law & Economics* **24**(3), 461–83.
URL: *http://ideas.repec.org/a/ucp/jlawec/v24y1981i3p461-83.html*

Guriev, S. and Treisman, D. (2015), How modern dictators survive: An informational theory of the new authoritarianism, Technical report, National Bureau of Economic Research.

Hallin, D. C. (1989), *The uncensored war: The media and Vietnam*, Univ of California Press.

Hellwig, C. and Veldkamp, L. (2009), 'Knowing what others know: Coordination motives in information acquisition', *The Review of Economic Studies* **76**(1), 223–251.

Hirshleifer, J. (1971), 'The private and social value of information and the reward to inventive activity', *The American Economic Review* pp. 561–574.

Ivanov, M. (2013), 'Information revelation in competitive markets', *Economic Theory* **52**(1), 337–365.

Janssen, M. C. and Roy, S. (2015), 'Competition, disclosure and signalling', *The Economic Journal* **125**, 86–114.

Kamenica, E. and Brad, L. E. (2014), 'Voters, dictators, and peons: expressive voting and pivotality', *Public Choice* **159**(1-2), 159–176.

Kamenica, E. and Gentzkow, M. (2011), 'Bayesian persuasion', *American Economic Review* **101**, 2590–2615.

Keller, G., Rady, S. and Cripps, M. (2005), 'Strategic experimentation with exponential bandits', *Econometrica* **73**(1), 39–68.

Kern, H. L. and Hainmueller, J. (2009), 'Opium for the masses: How foreign media can stabilize authoritarian regimes', *Political Analysis* **17**(4), 377–399.
URL: *http://pan.oxfordjournals.org/content/17/4/377.abstract*

Kim, J. and Fey, M. (2007), 'The swing voter's curse with adversarial preferences', *Journal of Economic Theory* **135**(1), 236–252.

Koessler, F. (2003), 'Persuasion games with higher-order uncertainty', *Journal of Economic Theory* **110**(2), 393–399.
URL: *http://ideas.repec.org/a/eee/jetheo/v110y2003i2p393-399.html*

Kolotilin, A. (2014), 'Optimal information disclosure: Quantity vs. quality', *UNSW Australian School of Business Research Paper* (2013-19).

Kris, E. (1941), 'German censorship instructions for the czech press', *Social Research* **8**(2), pp. 238–246.
URL: *http://www.jstor.org/stable/40981774*

Lawson, C. and McCann, J. A. (2005), 'Television news, mexico's 2000 elections and media effects in emerging democracies', *British Journal of Political Science* **35**(01), 1–30.
URL: *http://dx.doi.org/10.1017/S0007123405000013*

Levin, D., Peck, J. and Ye, L. (2005), 'Quality disclosure and competition', *Journal of Industrial Economics, forthcoming* .

Lipset, S. M. (1959), 'Some social requisites of democracy: Economic development and political legitimacy', *The American Political Science Review* **53**(1), pp. 69–105.
URL: *http://www.jstor.org/stable/1951731*

Lizzeri, A. and Yariv, L. (2013), 'Sequential deliberation', **Mimeo**.

Mannix, E. and Neale, M. A. (2005), 'What differences make a difference? the promise and reality of diverse teams in organizations', *Psychological science in the public interest* **6**(2), 31–55.

Mathis, J. (2008), 'Full revelation of information in sender-receiver games of persuasion', *Journal of Economic Theory* **143**(1), 571–584.
URL: *http://ideas.repec.org/a/eee/jetheo/v143y2008i1p571-584.html*

Mauro, P. (1995), 'Corruption and growth', *The quarterly journal of economics* pp. 681–712.

Menezes, F. M. and Monteiro, P. K. (2000), 'Auctions with endogenous participation', *Review of Economic Design* **5**(1), 71–89.

Messner, M. and Polborn, M. K. (2012), 'The option to wait in collective decisions and optimal majority rules', *Journal of Public Economics* **96**(5), 524–540.

Milgrom, P. R. (1981), 'Good news and bad news: Representation theorems and applications', *Bell Journal of Economics* **12**(2), 380–391.
URL: *http://ideas.repec.org/a/rje/bellje/v12y1981iautumnp380-391.html*

Moldovanu, B. and Shi, X. (2013), 'Specialization and partisanship in committee search', *Theoretical Economics* **8**(3), 751–774.
URL: *http://dx.doi.org/10.3982/TE1292*

Montalvo, J. G. and Reynal-Querol, M. (2005), 'Ethnic polarization, potential conflict, and civil wars', *American economic review* pp. 796–816.

Moreno, D. and Wooders, J. (2011), 'Auctions with heterogeneous entry costs', *The RAND Journal of Economics* **42**(2), 313–336.

Myatt, D. P. and Wallace, C. (2012), 'Endogenous information acquisition in coordination games', *The Review of Economic Studies* **79**(1), 340–374.

Persico, N. (2004), 'Committee design with endogenous information', *The Review of Economic Studies* **71**(1), 165–191.

Rayo, L. and Segal, I. (2010), 'Optimal information disclosure', *Journal of Political Economy* **118**(5), 949 – 987.
URL: *http://ideas.repec.org/a/ucp/jpolec/doi10.1086-657922.html*

Rodrik, D. (1993), 'The positive economics of policy reform', *The American Economic Review* pp. 356–361.

Seidmann, D. J. and Winter, E. (1997), 'Strategic information transmission with verifiable messages', *Econometrica* **65**(1), 163–170.
URL: *http://ideas.repec.org/a/ecm/emetrp/v65y1997i1p163-170.html*

Shadmehr, M. and Bernhardt, D. (2015), 'State censorship', *American Economic Journal: Microeconomics* **7**(2), 280–307.
URL: *http://www.aeaweb.org/articles.php?doi=10.1257/mic.20130221*

Shin, H. S. (1994), 'The burden of proof in a game of persuasion', *Journal of Economic Theory* **64**(1), 253–264.
URL: *http://ideas.repec.org/a/eee/jetheo/v64y1994i1p253-264.html*

Spence, M. (1973), 'Job market signaling', *The quarterly journal of Economics* pp. 355–374.

Stegeman, M. (1996), 'Participation costs and efficient auctions', *Journal of Economic Theory* **71**(1), 228–259.

Stivers, A. E. (2004), 'Unraveling of information: Competition and uncertainty', *Topics in Theoretical Economics* **4**(1).

Strulovici, B. (2010), 'Learning while voting: Determinants of collective experimentation', *Econometrica* **78**(3), 933–971.

Visser, B. and Swank, O. H. (2007), 'On committees of experts', *The Quarterly Journal of Economics* **122**(1), 337–372.

Wolinsky, A. (2003), 'Information transmission when the sender's preferences are uncertain', *Games and Economic Behavior* **42**(2), 319–326.
URL: *http://ideas.repec.org/a/eee/gamebe/v42y2003i2p319-326.html*

Zolotov, A. J. (2002), 'The dangers of riding the escalator', *The Moscow Times* (2417).