Essays on Information and Political Economy

Ali Simsek

A dissertation

submitted to the Faculty of

the department of Economics

in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy.

Boston College
Morrissey College of Arts and Sciences
Graduate School

March 2023

Essays on Information and Political Economy

Ali Simsek

Advisor: Mehmet Ekmekci, Ph.D.

Abstract

This dissertation consists of three essays on media, political and learning. More specifically, I investigate the effects of biased media and learning from that biased media on political institutions.

In the first essay, titled "Optimal Dynamic Information Supply and Competition", I provide a model of an information market where the viewers acquire signals each period at an attention cost, solving an optimal stopping problem à la Wald (1947), and the objective of the potentially biased information providers is to maximize the number of viewers who acquire signals from them across periods. I find that, in a monopoly market, the information provider sends unbiased signals that perfectly reveal the state of the world when there is a single period but provides biased signals when there are multiple periods. This is because biased signals elongate the learning process of some viewers, potentially increasing the information provider payoff. I also find that incentives due to competition, modeled as another information provider that is potentially biased in the opposite direction, overtake the intertemporal incentives and the full information equilibrium is recovered, even though it is wasteful in terms of social welfare. Hence, the paper provides a model with rational information providers and viewers that leads to biased signals in equilibrium.

In the second essay, titled "Voter Behavior and Information Aggregation in Elections with Supermajority", I provide a model of elections where there are three possible outcomes, but the voters can directly vote for one of the two options. The

outcome of the election corresponds to the options if the vote share for one of them is higher than a supermajority threshold. If neither of the options achieves that, then the result is the third outcome that the voters cannot explicitly vote for, which I interpret as compromise. I investigate various properties of elections in this setting. I find that, in line with the popular argument, supermajority rules foster compromise outcomes. But, on the other hand, elections with supermajority rules fail to aggregate information.

In the third essay, titled “Protests, Strategic Information Provision and Political Communication”, I consider a model of protests where the protesters learn about the state of the world via a biased information provider whose objective is to either instigate or dissuade the protest. A successful protest removes the incumbent from office, where the success threshold is determined by the incumbent who is biased. My main aim is to uncover whether the incumbent can learn the true state of the world from the protest turnout, even though the information of the citizens is provided by biased media. I pin down the optimal success threshold and signal noise choices by the incumbent and the information provider, respectively. I find that if the information provider is trying to instigate the protest, then political communication is always possible, regardless of the level of the bias of the incumbent. If the information provider is trying to dissuade the protest, then political communication is possible if and only if the incumbent bias is relatively small.

Contents

List of Figures

Chapter 1

Optimal Dynamic Information Supply and Competition

1.1 Introduction

Many empirical studies across multiple disciplines documented the existence of media bias (Puglisi & Snyder Jr, 2015; D'Alessio & Allen, 2000; Groeling, 2013). There are various explanations as to why that is the case, exploring incentives of information providers to provide biased information (Gentzkow et al., 2015). Some supply side arguments involve external incentives to the information providers from third parties, such as governments or advertisers (Prat, 2015). Furthermore, there are other demand side arguments that propose viewers with confirmatory bias as the reason, in other words, agents who derive higher utility from receiving information that confirm their prior beliefs. Hence, many explanations rely on external sources of incentives or agents with 'behavioral anomalies' in the sense of neoclassical economic modeling.

In this paper, we outline an environment with no such assumptions, and demonstrate that media bias can occur via the interaction of a profit maximizing information provider and rational viewers. We provide a model where the intertemporal incentives of the information provider lead to biased information in equilibrium. Furthermore, we investigate how competition affects this result.

We consider a model where the viewers (the agents who seek information) have to make a decision in the future, and they have the option to acquire signals from an information provider. They can receive a signal at any given period, or they can stop learning and make a decision. Importantly, they incur an attention cost for each signal they receive, hence learning is costly[1]. This is similar to the learning environment that is introduced by the influential paper by Wald (1947). Our formulation of the viewers side of the information market closely resembles Che & Mierendorff (2019).

The information providers are assumed to maximize the number of viewers who acquire signals from them across periods. They are also potentially biased. For a given state of the world x, if an information provider is x-biased, then it is truthful in state x, i.e., it sends signals which says that the state is x if the true state is indeed x. However, if the state is not x, there is still a potentially positive probability that the information provider lies and sends signals which says the state is x regardless. One implication of such a formulation is that upon observing a signal that says the state is not x by an x-biased information provider, the viewers know for sure that the state is not x. This type of learning environments with 'revelatory signals' is well suited to study political or scientific discovery processes, as well as criminal proceedings. The information providers choose the accuracy of their signals. Hence, they can choose to send unbiased or biased signals, as well as the amount of bias. The cost of production is an increasing function of the accuracy of the signals. This is because biased signals are less accurate, which we assume to be cheaper to produce.

Building on these preliminaries, we provide multiple versions of this model in terms of time horizon and level of competition. First, we start with the monopoly model with one period. We find that, under relevant parametric restrictions, the information provider sends fully informative signals, i.e. it has no bias. We then

[1]A model where the viewers buy signals from the information providers would lead to the same results qualitatively.

move on to a monopoly model with two periods. In one of the most important results of this paper, we find that the information provider does not send fully informative signals, hence disseminates biased information. This result is simply due to the intertemporal incentives of the information provider. If it were to send perfectly informative signals and fully reveal the state of the world, it would have no viewers in the second period. Hence, the information provider commits to sending biased signals in the hopes that some of the viewers continue to acquire information for the second period as well.

We then move on to our investigation of the effects of competition on our findings above. This is important because competition potentially incentivizes the information providers to send less biased signals, but also decreases the potential set of viewers. We find that the full information result still holds in a one-period market with two opposite-state-biased information providers. The competition, even though it eats away from their payoffs, is not enough to nudge the providers towards less accurate signals. As another important result of this paper, for a competitive model with multiple periods, we find that the full information equilibrium is recovered. More specifically, both information providers send unbiased signals in an equilibrium of the model, which is also the unique fully informative equilibrium. Hence, in contrast with the result for the one-period monopoly model, the incentives due competition overtakes the intertemporal incentives of the information providers. However, notice that this equilibrium is wasteful in terms of welfare of the market participants, since there is no need for two information providers incurring the cost of producing fully informative signals.

The rest of the paper is structured as follows. Section 2 goes over the relevant literature. Section 3 introduces the common preliminaries of the models. Section 4 explores the models with single information providers, in both one-period and two-period settings. Section 5 moves to the models with multiple information providers, again in both one-period and two-period settings. Section 6 contains concluding remarks. All proofs are relegated to the Appendix.

1.2 Related Literature

This paper builds on and contributes to a few different strands of literature. First and foremost, the viewer side of the market is formalized as a mass of individuals who are facing a version of the celebrated and influential 'optimal stopping problem', as introduced by Wald (1947) and further developed by Arrow et al. (1949). In these problems, decision makers face a usually binary choice under imperfect information, and at any time they have the option to either receive further information or stop and make a decision (hence the name). In this paper, in a similar vein, the viewer side of the market is modeled as decision makers who can, at any time, acquire a signal from an information source or make a decision and leave the market without acquiring a signal. This setup for the viewers is closely related to Che & Mierendorff (2019), as it is essentially a discrete time version of their model. Obviously this paper differs from the aforementioned paper as this one incorporates strategic information sources.

This paper is also closely related to a growing number of papers that concerns themselves with the decisions of the information providers under similar environments. Arguably the most similar one is Galperti & Trevino (2020), where they consider a competitive market of information supply. Similar to our formulation, the key driver of revenue for the information sources is the attention they receive, so their main motive is to maximize the number viewers who get information from them. But their formulation of the viewer side of the market, which builds on Dewan & Myatt (2008), Dewan & Myatt (2012) and Myatt & Wallace (2012), is completely different to ours. Namely, their viewers do not solve an optimal stopping problem (and hence the entire learning process is different). Even though some of the results are similar, especially the ones regarding the effects of competition, we uncover some novel results, such as the effects of the learning dynamics on the information source under the multi-period monopoly models. Another similar work is Lipnowski et al. (2020), where they consider the problem of a benevolent information provider who does not internalize the attention cost of the viewer. Clearly, the facts that the information source has the same motives as the viewer, that it is ignorant of the viewer's attention cost and that it is a monopoly makes this work completely distinct to ours. Lastly, another similar work to ours is Burke (2008), where they consider the dynamic problem of a possibly information source that tries to maximize its viewership (hence the same problem as ours) given that the viewers have confirmatory bias, i.e. they prefer information that reinforce their current beliefs. This formulation of the viewers is clearly different than ours, since the viewers in our case exhibits both confirmatory and contradictory learning behavior.

Furthermore, this paper contributes to the growing literature regarding learning under biased sources by incorporating strategic sources. The literature on optimal stopping problems with endogenous choice of information acquisition that focusses on Poisson processes for the signals (like our case) has been developing after the extensive focus on the drift diffusion models. The latter literature, which assumes that the signals follow a Brownian motion with an exogenous constant drift and the decision maker chooses the 'intensity' of her learning, consists of many important papers such as Moscarini & Smith (2001), Chan et al. (2018), Fudenberg et al. (2018), Ke & Villas-Boas (2019) and Henry & Ottaviani (2019). The former literature, which assumes that the decision makers choose how much attention to give to possibly biased information sources, is more open for future research, and the seminal works include the aforementioned Che & Mierendorff (2019), along with Nikandrova & Pancs (2018) and Mayskaya (2022). As Che & Mierendorff (2019) points out, Poisson processes are more suitable for modeling learning processes that involve discontinuous belief updating that mimics discoveries in real life, similar to processes regarding criminal proceedings, scientific breakthroughs and political information. The last example constitutes the basis of our modeling foundations, even though the model can be applied to any such learning environment.

This paper and the learning under biased information sources literature in general are closely related to the highly influential rational inattention literature that is introduced by Sims (2003). Similar to that model, the viewer side of the market in this paper chooses which source to learn from (under the formulations with multiple sources), or to learn at all. This action space closely resembles rational inattention models. But the learning process that is used in the modeling of this paper implies that there is a constant marginal cost to learning, which is different than the rational inattention models.

Finally, this paper also contributes to the larger literature regarding the political economy of mass media that spans various disciplines along with economics. An excellent review on the literature, which is somewhat dated at this time, is presented in Prat & Strömberg (2013).

1.3 The Model

There is a unit measure set of viewers, denoted by $i \in V = [0,1]$, and they each have to take an irreversible action $a \in A = \{l, r\}$ by the end of the game. The payoff from this action, denoted by u_a^ω, depends on the unknown state of the world $\omega \in \{L, R\}$. If the action matches the state of the world, it yields utility of 1, and -1 otherwise. For each $i \in V$, her belief at the end of period t that the state is R is denoted by ρ_i^t. Initially, $\rho_i^0 = i$ for all $i \in V$. In other words, we assume that the priors of the viewers are distributed uniformly on $[0,1]$ and a viewer's identity is equal to her initial belief. Naturally, if the viewer decides to take an action given her current belief, the optimal action is determined by

$$a^*(\rho_i^t) = \begin{cases} l, & \text{if } \rho_i^t < \frac{1}{2} \\ r, & \text{if } \rho_i^t > \frac{1}{2} \\ \triangle(A), & \text{if } \rho_i^t = \frac{1}{2} \end{cases}$$

There are T periods, and each viewer can take the irreversible action at the end of any period in the game or before the game starts. Otherwise, they can decide to acquire information from an information source. If they decide to do so, they get a signal s^I from information source I. We assume that viewers can pay attention to at most one information source[2]. Notice that they have to either make a decision and leave the game or acquire information from a source. Hence, the viewers are facing a stopping problem. For any period without a decision made, each viewer incurs an attention cost $c > 0$, since she acquires information. For simplicity, we assume that $c \in \left(0, \frac{1}{2}\right)$, otherwise the attention cost would be prohibitively costly in the two-period models. We take this as the benchmark, and assume that the cost of acquisition is low enough so that acquiring a signal from an information provider who is fully informative in one state and fully uninformative in the other is feasible.

An information source I can be either R-biased or L-biased and they each send a signal $s^I \in \{s^I_a\}_{a\in\{r,l\}}$. At the beginning of the game, each information source commits to a parameter $\lambda \in [0, 1]$ that determines a probability distribution for its signals, which is stationary across periods. These parameters are common knowledge. If an information source is L-biased then its signal structure is given by

State/Signal	s^L_l	s^L_r
L	1	0
R	$1 - \lambda_L$	λ_L

and if an information source is R-biased then its signal structure is given by

State/Signal	s^R_l	s^R_r
L	λ_R	$1 - \lambda_R$
R	0	1

An L-biased information source always tells the truth when the state of the world is L and reveals the true state. But when the state is R, it tells the truth

[2]This assumption does not affect our results since there are infinitely many viewers and only finitely many viewers can an incentive to 'multi-home' in our model, hence their effect on the information provider payoff is measure-zero.

and reveals that the state is R only with probability $\lambda_L \in [0,1]$, and lies and says that the state of the world is L with the remaining probability $1-\lambda_L$ (hence the name L-biased). Similarly, an R-biased information source always tells the truth when the state of the world is R and reveals the true state. But when the state is L, it tells the truth and reveals that the state is L only with probability $\lambda_R \in [0,1]$, and lies and says that the state of the world is R with the remaining probability $1-\lambda_R$. Notice that this formulation has the following important consequence: If an L-biased information source sends the signal s_r^L, then the state is R with probability 1, and if an R-biased information source sends the signal s_l^R, then the state is L with probability 1. But the other signals do not perfectly reveal the state, since there is strictly positive probability that an L-biased information source sends s_l^L in both states, and similarly, there is strictly positive probability that an R-biased information source sends s_r^R in both states. Hence, in effect, λ_L and λ_R parametrize the informativeness of the signals that are coming from an L-biased source and an R-biased source, respectively.

The information sources also hold initial beliefs regarding the state of the world, denoted by P which is their belief that the state of the world is R. We normalize $P=\frac{1}{2}$, which is equal to the average viewer belief. Each information source I maximizes the expected measure of viewers who acquire information from it across time. But the information production is costly and the cost depends on λ. We assume that the per-period cost of information production is $\Gamma(\lambda)=a\lambda^2$ for some $a \in \left(0, \frac{1-c}{4}\right)$, a simple convex function. As a result, sending completely uninformative signals (an L-biased information source always sending the signal s_l^L and an R-biased information source always sending the signal s_r^R) is costless and sending completely informative signals that perfectly reveal the state of the world has the costs a. Furthermore, we assume that $a \in \left(0, \frac{1-c}{4}\right)$ to make sure that sending perfectly informative signals is not prohibitively costly in any of the versions of the model presented in this paper. More specifically, we assume that sending a perfectly informative signal yields a higher payoff to the information provider than sending a completely uninformative signal. This way, if we find any noise in the signal(s) at the optimum, we know that is it either due to the effects of multiple periods, or competition or both. For simplicity, we assume that neither the viewers nor the information sources discount future payoffs. We normalize the price of attention to 1. Hence, this cost function can be though as a ratio of the cost to the total value of the market. We also assume that the signal is still produced even if no viewer decides to take it, so the cost of information production still applies.

In sum, the timing of the game is as follows: Nature draws a state of the world, unknown to all players. Then, information sources in the market simultaneously decide on their respective λ's, given their priors, and these λ's become publicly known. Then, in period 1, each viewer decides on taking an action now, which is costless, and exiting the game or receiving a signal from an information source at an attention cost c and make a decision after. Then in period $t \neq 1$, the timing is exactly the same as the first period, but only the viewers who did not previously make a decision can acquire information and take a decision. Then, at the end of period T, the state of the world is revealed and payoffs are realized.

Next, we characterize different versions of the problem with various number of information sources and periods. Our main aims are to understand the effects of multiple periods as opposed to one period, and the effect of multiple information sources as opposed to one. For that reason, we start with the simplest problem where there is a single period and a single information source as the benchmark, then we gradually add more periods and information sources. The equilibrium notion we employ is the perfect Bayesian equilibrium. We assume that, when indifferent, viewers choose to acquire the information rather than not, and acquire information with equal probability when they are indifferent between multiple information sources. Moreover, we also assume that the information source provides the signal with the highest informativeness when it is indifferent.

1.4 Single Information Provider

In this section, we assume that there is a monopoly information provider who is, without loss of generality, R-biased. In the immediate subsection below, we assume $T = 1$, i.e. there is a single period, hence the viewers can acquire a signal only once. In the following subsection, we assume $T = 2$, so the viewers can acquire information twice. Our main aim in this section is to lay out the intertemporal incentives of the information provider (hence the focus on the single provider case). In order to achieve our aim, we pin down the equilibrium signal precision levels in these two cases and compare them, reveal when they are equal to each other and when they are not.

1.4.1 One-Period Problem with One Information Source

In this first subsection, we solve the static problem with one information source. If a viewer i does not acquire information, the payoff from taking an optimal action is

$$u^0(i) = \max\{i1 + (1-i)(-1), i(-1) + (1-i)1\} = \max\{2i-1, 1-2i\}$$

where the first term in the maximum function is the expected utility from taking the action r and the second term is the expected utility from taking the action l. Furthermore, if the viewer decides to acquire information, then her expected utility at the beginning of the game is

$$u^1_R(i, c, \lambda_R) = [i + (1-i)(1-\lambda_R)] \max\left\{\tfrac{2i}{1-\lambda_R+\lambda_R i} - 1, 1 - \tfrac{2i}{1-\lambda_R+\lambda_R i}\right\} + (1-i)\lambda_R - c$$

where the first term in the summation is the expected utility of the viewer when there is no conclusive signal from the information source, i.e., when the R-biased information source sends s^R_r. This happens when either $\omega = R$ (which the viewer believes that it happens with probability i according to her prior belief) and the

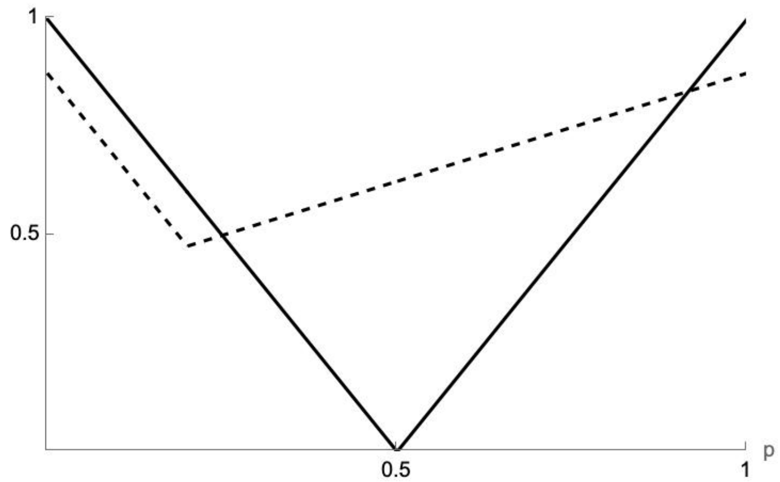

Figure 1.1: The figure plots $u^0(i)$ (solid) and $u^1_R(i, c, \lambda_R)$ (dashed) when $\lambda_R = 0.75$ and $c = 0.125$.

information source reveals the true state with probability 1, or $\omega = L$ (which the viewer believes that it happens with probability $1 - i$) but the information source sends the "incorrect" signal with probability $1 - \lambda_R$. In such a case, the viewer updates her initial belief according to the Bayes' rule and calculates the expected utilities from available actions r and l, and chooses the action that yields the highest expected utility according to her updated belief. The second term in the summation signifies the expected utility of the viewer then she receives a conclusive signal from the information source, i.e., when the R-biased information source sends s^R_l. This only happens when $\omega = L$ with probability λ_R. In such a case, the viewer knows that the state is L, hence takes the action l and receives the payoff of 1. Lastly, since the viewer acquired information, she incurs the attention cost c.

Given c and λ_R, we can find the cutoffs for the sets of viewers who acquire information and who immediately take an action. At these cutoff beliefs, the viewers will be indifferent between acquiring information and taking an action immediately. Denote the set of viewers who choose to acquire information from the R-biased information source with $\mathcal{A}^R \subseteq [0, 1]$. A few steps of algebra show that

$$\mathcal{A}^R = \begin{cases} \left[\frac{2(1-\lambda_R)+c}{2(2-\lambda_R)}, 1 - \frac{c}{2\lambda_R}\right], & \text{if } \lambda_R \geq c \\ \emptyset, & \text{otherwise} \end{cases}$$

We sometimes refer to $\mathcal{A}^R$ simply as the attention set. An important note is that $\mathcal{A}^R$ is convex valued. This implies that the information source cannot cater to extreme ends of the belief spectrum at the same time. Furthermore, there always exists viewers who do not acquire information and directly make a decision, and these viewers are concentrated in the extreme edges of the prior belief spectrum.

As explained earlier, the information source's objective is to maximize the mass of citizens acquiring information from it, denoted by $|\mathcal{A}^R|$, minus the cost of information. In order to better understand the problem of the information provider, we first examine the effects of the model parameters on the cardinality of the attention set, provided in the lemma below.

Lemma 1.4.1. *(i)* $|\mathcal{A}^R|$ *is increasing in* λ_R.

(ii) $|\mathcal{A}^R|$ *is decreasing in* c.

As expected, the attention set is larger when the signals are more informative. This is not surprising because more informative signals generate higher expected utilities for the viewers, especially the ones that have intermediary priors that are further away from 0 or 1, hence more viewers find it beneficial to acquire the signals at the associated cost. Similarly, given the informativeness, a higher cost of signal acquisition leads to a smaller attention set, since fewer viewers find it beneficial at a higher cost.

Equipped with a better understanding of the attention set, we now turn to the problem of the information source. As explained above, the information source chooses λ_R to maximize $|\mathcal{A}^R|$ minus the resulting cost. More specifically, the payoff to the information source for a specific value of λ_R given c is

$$\pi^R = \begin{cases} 1 - \frac{c}{2\lambda_R} - \frac{2(1-\lambda_R)+c}{2(2-\lambda_R)} - a\lambda_R^2, & \text{if } c \leq \lambda_R \leq 1 \\ -a\lambda_R^2, & \text{otherwise} \end{cases}$$

The following result provides the optimal informativeness of the signals sent by the information source.

Theorem 1.4.2. $\lambda_R^* = 1$ *for all* c *and* a*.*

The result above demonstrates that the information provider would send perfectly informative signals for any values of $c \in \left(0, \frac{1}{2}\right)$ and $a \in \left(0, \frac{1-c}{4}\right)$. Hence, the information provider has no incentive to decrease the informativeness of the signal in order to decrease the cost of information production, in the absence of any other incentives that might stem from multiple periods or competition. This obviously depends on our assumption regarding the value of a. If it was prohibitively high, we would not be able to reach to this conclusion. But taking this as the benchmark, in the next subsections of the paper, we try to understand how other incentives, added by either multiple periods or the existence of competition or both affect the value of the optimal informativeness.

1.4.2 Two-Period Problem with One Information Source

In this subsection, we investigate a model with two periods. The setup is exactly the same, except that the viewers can acquire information twice, and the information source tries to maximize the expected attention it gets across periods. We assume that the prior of the information source regarding the state of the world is equal to the average viewer prior, which is $\frac{1}{2}$.

The expected utilities of viewer i when she decides to acquire no information and acquire information only in the first period are the same as above. If the viewer decides to acquire information for two periods, then her expected utility is

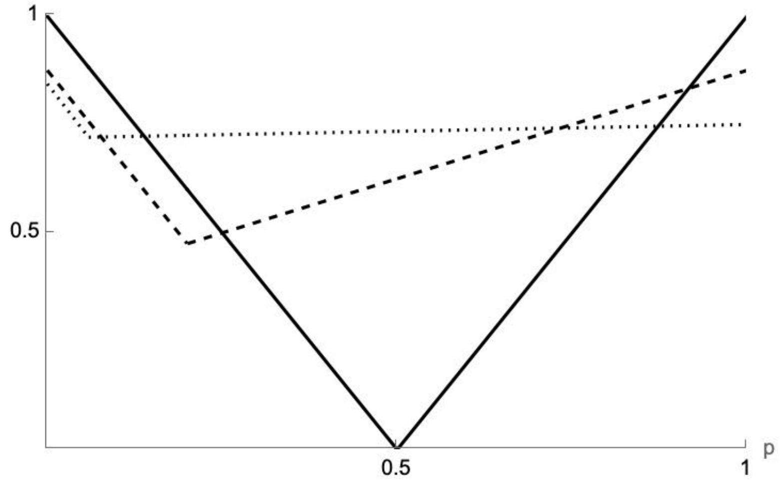

Figure 1.2: The figure plots $u^0(i)$ (solid) and $u_R^1(i, c, \lambda_R)$ (dashed) and $u_R^2(i, c, \lambda_R)$ (dotted) when $\lambda_R = 0.75$ and $c = 0.125$.

given by

$$
\begin{aligned}
u_R^2(i, c, \lambda_R) =& [i + (1-i)(1-\lambda_R)^2] \\
& \left(\max\left\{1 - \frac{2i}{1-(1-i)(2-\lambda_R)\lambda_R}, \frac{2i}{1-(1-i)(2-\lambda_R)\lambda_R} - 1\right\} - 2c \right) \\
& + (1-i)\lambda_R(1-c) + (1-i)(1-\lambda_R)\lambda_R(1-2c)
\end{aligned}
$$

where the first term is her expected utility when she does not receive a conclusive signal, the second term is her expected utility when she receives a conclusive signal in the first period and the last term is her expected utility when she receives a conclusive signal in the second period (as noted above, the conclusive signal is s_l^R and the non-conclusive signal is s_r^R). Given this, the set of viewers who acquire information at each period can be obtained. Denote the set of viewers who choose to acquire information from the R-biased information source in period 1 and period 2 (conditional on not observing the conclusive signal) by $\mathcal{A}_1^R$ and $\mathcal{A}_2^R$, respectively. A few steps of algebra reveal the following boundaries for these sets.

$$\mathcal{A}_1^R = \begin{cases} \left[\frac{2(1-\lambda_R)+c}{2(2-\lambda_R)}, 1 - \frac{c}{2\lambda_R}\right], & \text{if } c < \widehat{c} \text{ and } c \leq \lambda_R < \underline{\lambda}_R \\ & \text{or } c < \widehat{c} \text{ and } \overline{\lambda}_R \leq \lambda_R \leq 1 \\ & \text{or } c \geq \widehat{c} \text{ and } c \leq \lambda_R \leq 1 \\ \left[1 - \frac{2(1-c)}{4-(4+c-2\lambda_R)\lambda_R}, 1 - \frac{c}{2\lambda_R}\right], & \text{if } c < \widehat{c} \text{ and } \underline{\lambda}_R \leq \lambda_R < \overline{\lambda}_R \\ \emptyset, & \text{otherwise} \end{cases}$$

$$\mathcal{A}_2^R = \begin{cases} \left[1 - \frac{2(1-c)}{4-(4+c-2\lambda_R)\lambda_R}, \frac{(2\lambda_R-c)(1-\lambda_R)}{(2-2\lambda_R+c)\lambda_R}\right], & \text{if } c \leq \widehat{c} \text{ and } \underline{\lambda}_R \leq \lambda_R < \overline{\lambda}_R \\ \emptyset, & \text{otherwise} \end{cases}$$

where $\widehat{c} = 1 - \sqrt{8\sqrt{2} - 11} \approx 0.44$, $\underline{\lambda}_R = \frac{1}{4}\left(\frac{4}{2-c} + c - \sqrt{c^2 + \frac{8(2-3(2-c)c)}{(2-c)^2}}\right)$ and $\overline{\lambda}_R = \frac{1}{4}\left(\frac{4}{2-c} + c + \sqrt{c^2 + \frac{8(2-3(2-c)c)}{(2-c)^2}}\right)$.

The attention sets for the first and second period reveal some interesting features of the model. If the attention cost is relatively high, then no viewer finds it profitable to acquire a signal in both periods, since it is too expensive in terms of attention cost. Hence, there exists a cutoff for the attention cost, denoted by $\widehat{c}$, such that if the attention cost is higher than $\widehat{c}$, the problem of the information provider is the same as the one period version, since no viewer acquires information, regardless of the value of λ_R. Similarly, if the informativeness of the signal is too low or too high, again no viewer acquires information in the second period. To see why the first part of the observation is true, consider the limit case where the signal is perfectly informative. In such a case, the signal would reveal the state perfectly in the first period, hence there would be no reason to acquire information in the second period. For the second part of the observation, if the informational content of the signal is relatively low, then a small portion of the viewers with beliefs that are close to the border beliefs would acquire information for one period only, with the expectation that the signal would confirm their belief and they cross the border belief to confidently vote r. However, if the precision of the signal is at an intermediary level, then some viewers find it beneficial to acquire information for two periods, since the benefit (a more extreme belief if there is no revelatory signal and knowing the state for sure if there is a revelatory signal) outweighs the cost of acquiring the signal twice.

Again, in order to better understand the incentives of the information provider, we first examine the effects of the model parameters on the cardinality of the period 1 and period 2 attention sets, provided in the lemma below.

Lemma 1.4.3. *(i)* $|\mathcal{A}_1^R|$ *is decreasing in* c.

(ii) $|\mathcal{A}_1^R|$ *is increasing in* λ_R.

(iii) $|\mathcal{A}_2^R|$ *is decreasing in* c.

(iv) $|\mathcal{A}_2^R|$ *is concave in* λ_R.

The lemma above reveals an interesting trade off regarding the signal precision parameter. As expected, the measure of viewers in both periods is decreasing in the attention cost. This is simply due to the fact that as the cost of attention increases, fewer and fewer viewers are going to find it beneficial to acquire information. Furthermore, similar to the result in Lemma 1.4.1, as the signals get more informative, the viewers who acquire information during period 1 increases. But, this is not true for the second period. In period 2, the mass of viewers who acquire information increases as the precision of the signal increases starting from low precision. This is because more and more viewers find it worthwhile to acquire information for two periods as precision increases, since higher precision makes it more likely to get a revelatory signal and also increases the belief update in the case where there is none. But as the signal gets more and more precise, the reverse happens; the mass of viewers who acquire information decreases as a result. When the precision is high enough, acquiring a signal for only one period actually yields a higher payoff than acquiring signals in both periods. To see why, consider the problem of a viewer who received a non-revelatory signal in period 1 and is now considering acquiring information again. Since the signal was s_r^R and λ_R was relatively high, she is now 'highly convinced' that the state is R, i.e., ρ_i^1

is relatively large. As discussed above as well, acquiring a signal has two benefits: the possibility of a revelatory signal and the update in the belief when there is no revelatory signal. Since the viewer's belief is relatively high, the belief jump from a possible non-revelatory signal is low. Furthermore, since she is convinced that the state is R, the possibility of a revelatory signal is very low from her perspective. These two effects, when combined, yield an expected utility to the viewer that does not compensate the attention cost associated with receiving an additional signal. As a result, when λ_R is relatively high and increases further, the mass of viewers who acquire a signal in the second period decreases.

Equipped with these results, we now turn to the problem of the information provider. Given c and the information source's belief, the expected payoff of the information source as a result of its choice of λ_R is pinned down by

$$\pi^R = \begin{cases} 1 - \frac{c}{2\lambda_R} - \frac{2(1-\lambda_R)+c}{2(2-\lambda_R)} - 2a\lambda_R^2, & \text{if } c < \widehat{c} \text{ and } c \leq \lambda_R < \underline{\lambda}_R \\ & \text{or } c < \widehat{c} \text{ and } \overline{\lambda}_R \leq \lambda_R \leq 1 \\ & \text{or } c \geq \widehat{c} \text{ and } c \leq \lambda_R \leq 1 \\ \left(2 - \frac{\lambda_R}{2}\right) \frac{2(1-c)}{4-(4+c-2\lambda_R)\lambda_R} - \frac{c}{2\lambda_R} & \\ + \left(1 - \frac{\lambda_R}{2}\right)\left(\frac{(2\lambda_R - c)(1-\lambda_R)}{(2-2\lambda_R+c)\lambda_R} - 1\right) - 2a\lambda_R^2, & \text{if } c < \widehat{c} \text{ and } \underline{\lambda}_R \leq \lambda_R < \overline{\lambda}_R \\ -2a\lambda_R^2, & \text{otherwise} \end{cases}$$

Given this expected payoff function, the following theorem pins down the optimal signal precision by the information provider when there are two periods.

Theorem 1.4.4. *There exists $\underline{c}_1$, $\overline{c}_1$ and $\widehat{a}_1$ such that*

(i) if $c < \underline{c}_1$ then $\lambda_R^* < 1$,

(ii) if $\underline{c}_1 < c < \overline{c}_1$ then $\lambda_R^* < 1$ if and only if $a > \widehat{a}_1$,

(iii) if $c > \overline{c}_1$ then $\lambda_R^* = 1$.

Theorem 1.4.4 reveals the effects of multiple periods on the incentives of the information provider. First assume that the attention cost is relatively low. In such a case, the information provider knows that, since the attention cost is low, some viewers find it beneficial to acquire signals for two periods, assuming that there is no revelatory signal in the first period. As a result, signal precision is determined by the following tradeoff: A higher precision signal increases the first period viewers and decreases the second period viewers, and a lower precision signal has the reverse effect. The result shows that the incentive due to the second period viewers is strong enough for the information provider that it finds it optimal so send noisy signals. Hence, the information provider sends biased signals in order to keep more viewers engaged for both periods. Now assume that the attention cost is at an intermediary level. Since the signals are more costly in this case, the incentive due to the second period viewers is lower. In such a case, the information provider sends perfectly informative signals if and only if the production cost of perfectly informative signals is low enough. Lastly, assume that the attention cost is relatively high. This implies that the number of two-period viewers is small to begin with, hence the incentive due to these viewers is not enough to convince the information provider to send noisy signals in order to keep the viewers engaged for two periods.

The following corollary juxtaposes the equilibria for the one-shot and two-period formulations of the game at hand. It directly follows from Theorems 1.4.2 and 1.4.4.

Corollary 1.4.4.1. λ_R^* *for the one period game is equal to the* λ_R^* *for the two-period game if and only if either* $\underline{c}_1 < c < \bar{c}_1$ *and* $a < \widehat{a}_1$*, or* $c > \bar{c}_1$*.*

The corollary above presents the comparison of the equilibrium signal precision for the R-biased information provider under the two formulations we presented up until this point: one period versus two periods, both with a monopoly information provider. In simple terms, the corollary says that the outcome of the one-shot game and the two-period game coincide (and lead to perfectly informative signals) under two distinct conditions: either the attention cost is neither high or low and the information production cost is relatively low, or the attention cost is relatively high (regardless of the value of the information production cost). Notice that, if the attention cost is relatively low, the equilibrium signal precisions never coincide, due to the fact that intertemporal incentives for the information provider are too large to disregard. This is because when the attention cost is relatively low, more viewers, concentrated in the middle of the prior belief distribution, are willing to receive signals for two periods rather than one. If the attention cost is at an intermediary level, like the first case in the corollary above, then the incentive to decrease the informativeness of the signal in order to gain more second period viewers still exists for the information provider, but prevails if and only if the information production cost is relatively large, which is another added incentive not to provide perfectly informative signals for the information provider. Hence, when the intertemporal incentive is relatively small, the lower cost of information production incentive must be present for the information provider to send perfectly informative signals. Lastly, when the attention cost relatively large, which leads to very few viewers being willing to acquire the signal in the second period, the intertemporal incentive is not enough for the information provider, hence it sends perfectly informative signals, no matter the cost associated with it.

1.5 Multiple Information Providers

In the previous subsection, we uncovered that, in the presence of intertemporal incentives, the R-biased information source provides fully informative signals if and only if the cost of attention is at an intermediate level and the cost of information production is relatively low or the cost of attention is at a relatively high level. In other words, if the attention cost is relatively low or it is at an intermediary level and the information production cost is relatively high, the information provider sends biased signals. This is in stark difference with the equilibrium for the one shot game where the information provider commits to perfectly revealing the state, as concluded in Corollary 1.4.4.1. In this section, we turn to investigating the

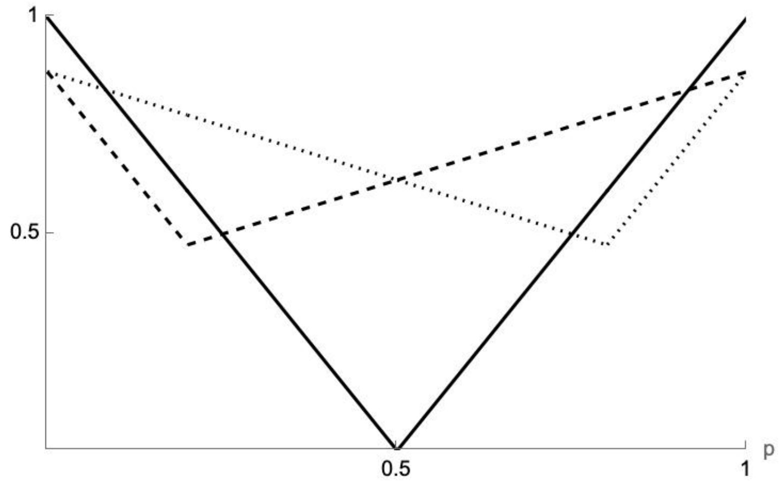

Figure 1.3: The figure plots $u^0(i)$ (solid) and $u^1_R(i, c, \lambda_R)$ (dashed) and $u^1_L(i, c, \lambda_L)$ (dotted) when $\lambda_R = 0.75$, $\lambda_L = 0.75$ and $c = 0.125$.

effects of competition in these models. In the following subsections, we add an L-biased information source to the market, pin down the equilibrium signal precisions in these markets, and compare them to their single source counterparts.

1.5.1 One-Period Problem with Two Information Sources

In subsection 1.4.1, we found that in a single period market, a monopoly information provider sends perfectly informative signals. We now investigate whether this result still holds when we add another information provider, more specifically a competing L-biased one. On one hand, the possible market size shrinks, since now some viewers find the competition more appealing, and along with the effects of the information production cost, this might lead to a decrease in the equilibrium signal precision. On the other hand, the incentives associated with competition might reinforce the equilibrium with fully informative signals.

For viewer i, the expected utility of not acquiring information, acquiring information once from the R-biased source and acquiring information from an R-biased source are given by $u^0(i)$ and $u^1_R(i, c, \lambda_R)$, respectively, in subsection 3.1. Similarly, the expected utility of viewer i from acquiring information once from the L-biased source is

$$u_L^1(i,c,\lambda_L) = [(1-i)+i(1-\lambda_L)]\max\left\{\frac{2(1-\lambda_L)i}{1-\lambda_L i}-1, 1-\frac{2(1-\lambda_L)i}{1-\lambda_L i}\right\}+i\lambda_L - c$$

where the first term in the summation is the expected utility of the viewer when there is no conclusive signal from the information source, i.e., when the L-biased information source sends s_l^L. This happens either when $\omega = L$ (which the viewer believes that it happens with probability $1-i$ according to her prior belief) and the information source reveals the true state with probability 1, or when $\omega = R$ (which the viewer believes that it happens with probability i) but the information source sends the "incorrect" signal with probability $1-\lambda_L$. In such a case, the viewer updates her initial belief according to the Bayes' rule and calculates the expected utilities from available actions r and l, and chooses the action that yields the highest expected utility according to her updated belief. The second term in the summation signifies the expected utility of the viewer then she receives a conclusive signal from the information source, i.e., when the L-biased information source sends s_r^L. This only happens when $\omega = R$ with probability λ_L. In such a case, the viewer knows that the state is R, hence takes the action r and receives the payoff of 1. Lastly, since the viewer acquired information, she incurs the attention cost c.

Given c, λ_R and λ_L, we can pin down the set of viewers who do not acquire a signal, acquire a signal from the R-biased information source and acquire a signal from the L-biased information source. Denote the set of viewers who choose to acquire information from the R-biased information source with $\mathcal{A}^R$ and the set of viewers who choose to acquire information from the L-biased information source with $\mathcal{A}^L$. A few steps of algebra show that

$$\mathcal{A}^R = \begin{cases} \left[\frac{2(1-\lambda_R)+c}{2(2-\lambda_R)}, 1 - \frac{c}{2\lambda_R}\right], & \text{if } \lambda_R \geq \frac{(2+c)\lambda_L - 2c}{2\lambda_L - c} \\ \left[\frac{1-\lambda_R}{2-\lambda_R-\lambda_L}, 1 - \frac{c}{2\lambda_R}\right], & \text{if } \frac{c(2-\lambda_L)}{2(1-\lambda_L)+c} \leq \lambda_R < \frac{(2+c)\lambda_L - 2c}{2\lambda_L - c} \\ \emptyset, & \text{if } \lambda_R < \frac{c(2-\lambda_L)}{2(1-\lambda_L)+c} \end{cases}$$

and

$$\mathcal{A}^L = \begin{cases} \left[\frac{c}{2\lambda_L}, \frac{2-c}{2(2-\lambda_L)}\right], & \text{if } \lambda_R \leq \frac{c(2-\lambda_L)}{2(1-\lambda_L)+c} \\ \left[\frac{c}{2\lambda_L}, \frac{1-\lambda_R}{2-\lambda_R-\lambda_L}\right], & \text{if } \frac{c(2-\lambda_L)}{2(1-\lambda_L)+c} < \lambda_R \leq \frac{(2+c)\lambda_L - 2c}{2\lambda_L - c} \\ \emptyset, & \text{if } \lambda_R > \frac{(2+c)\lambda_L - 2c}{2\lambda_L - c} \end{cases}$$

We sometimes refer these simply as the attention sets. Again, both of the attention sets are convex. This implies that the information sources cannot cater to extreme ends at the belief spectrum at the same time. Furthermore, if λ_R is relatively large compared to λ_L (when $\lambda_R > \frac{(2+c)\lambda_L - 2c}{2\lambda_L - c}$), the R-biased information source captures the whole market, i.e., the viewers either acquire information from the R-biased information source or they do not acquire information at all. If λ_L is relatively large compared to λ_R (when $\lambda_R < \frac{c(2-\lambda_L)}{2(1-\lambda_L)+c}$), the L-biased information source captures the whole market, i.e., the viewers either acquire information from the L-biased information source or they do not acquire information at all. If neither is true and one is not larger relative to the other (when $\frac{c(2-\lambda_L)}{2(1-\lambda_L)+c} < \lambda_R \leq \frac{(2+c)\lambda_L - 2c}{2\lambda_L - c}$), no information source fully captures the market and they both get some portion of the viewers. Similar to the monopoly information source case, there always exists viewers who do not acquire information and directly make a decision, and these viewers are concentrated in the extreme edges of the prior belief spectrum.

Again, the information sources' objective is to maximize the mass of citizens acquiring information from them, minus the cost of information production they incur. In order to better understand the problem of the information providers, we first examine the effects of the model parameters on the cardinality of the attention sets, provided in the lemma below.

Lemma 1.5.1. *(i) $|\mathcal{A}^R|$ is increasing in λ_R.*

(ii) $|\mathcal{A}^R|$ is decreasing in λ_L.

(iii) $|\mathcal{A}^R|$ is decreasing in c.

(iv) $|\mathcal{A}^L|$ is increasing in λ_L.

(v) $|\mathcal{A}^L|$ is decreasing in λ_R.

(vi) $|\mathcal{A}^L|$ is decreasing in c.

The lemma above shows that, as expected, the R-biased source's attention set is increasing in its own informativeness and decreasing in the attention cost, similar to the monopoly case. Furthermore, it is also decreasing in the L-biased information source's informativeness. This is expected since the more informative the L-biased information source's signals are, the more viewers find it beneficial to acquire information from the L-biased information source rather than the R-biased information source, leading to a shrinkage in the attention set of the R-biased information source. Similarly, the L-biased source's attention set is increasing in its own informativeness and decreasing in the attention cost and the R-biased source's informativeness.

Equipped with a better understanding of the properties of the attention sets, once again, we turn to the problem of the information sources. The payoffs to the information sources can be written as

$$\pi^R = \begin{cases} \frac{1}{2}\left(1 - \frac{c}{2\lambda_R} - \frac{2(1-\lambda_R)+c}{2(2-\lambda_R)}\right) - a\lambda_R^2, & \text{if } \lambda_R = \lambda_L = 1 \\ 1 - \frac{c}{2\lambda_R} - \frac{2(1-\lambda_R)+c}{2(2-\lambda_R)} - a\lambda_R^2, & \text{if } \lambda_R \geq \frac{(2+c)\lambda_L - 2c}{2\lambda_L - c} \\ 1 - \frac{c}{2\lambda_R} - \frac{1-\lambda_R}{2-\lambda_R-\lambda_L} - a\lambda_R^2, & \text{if } \frac{c(2-\lambda_L)}{2(1-\lambda_L)+c} \leq \lambda_R < \frac{(2+c)\lambda_L - 2c}{2\lambda_L - c} \\ -a\lambda_R^2, & \text{if } \lambda_R < \frac{c(2-\lambda_L)}{2(1-\lambda_L)+c} \end{cases}$$

and

$$\pi^L = \begin{cases} \frac{1}{2}\left(\frac{2-c}{2(2-\lambda_L)} - \frac{c}{2\lambda_L}\right) - a\lambda_L^2, & \text{if } \lambda_R = \lambda_L = 1 \\ \frac{2-c}{2(2-\lambda_L)} - \frac{c}{2\lambda_L} - a\lambda_L^2, & \text{if } \lambda_R \leq \frac{c(2-\lambda_L)}{2(1-\lambda_L)+c} \\ \frac{1-\lambda_R}{2-\lambda_R-\lambda_L} - \frac{c}{2\lambda_L} - a\lambda_L^2, & \text{if } \frac{c(2-\lambda_L)}{2(1-\lambda_L)+c} < \lambda_R \leq \frac{(2+c)\lambda_L-2c}{2\lambda_L-c} \\ -a\lambda_L^2, & \text{if } \lambda_R > \frac{(2+c)\lambda_L-2c}{2\lambda_L-c} \end{cases}$$

The following result provides the optimal informativeness of the signals sent by the information sources.

Theorem 1.5.2. *$\lambda_R^* = 1$ and $\lambda_L^* = 1$ is the unique equilibrium for all c and a.*

The result shows that, in equilibrium, both information sources are going to provide signals that perfectly reveal the state. This is simply due to the fact that both payoff functions are strictly increasing in their own informativeness (regardless of the competitor's informativeness and the values of the attention and information production costs) and yield positive payoff when the signals are perfectly informative. Remember we also found that, in the one period problem with a single information source, the provider sends perfectly informative signals as well. This result implies that the competition does not eliminate the incentives to provide full information, even though it shrinks the set of viewers and leads to lower payoff. The equilibrium payoff of the R-biased information source is $1-c-a$ under the one period-one information source case and goes down to $\frac{1-c}{2} - a$ under the one period-two information sources case.

Lastly, we turn to the effects of competition on welfare. We define welfare simply as the sum of the payoffs of the viewers and the information providers in the market. The result below compares the welfare of the viewers and the information sources under monopoly and competition.

Lemma 1.5.3. *The total welfare of the viewers and the information providers is higher in the equilibrium of the one period-one source model compared to the one period-two sources model.*

The lemma above shows that, for the one shot models, the total market payoff under monopoly is higher than the total market payoff under two information sources. In other words, competition leads to welfare loss. More specifically, the competition leads to no change in the total viewer welfare (since the same viewers receive perfectly informative signals in both models) but it leads to a decrease in the total provider welfare. This is because the total mass of viewers is the same (hence the total market revenue is the same) but two providers incur the cost of producing perfectly informative signals as opposed to one. This additional information production cost, which does not correspond to an increase in revenue, leads to a welfare loss for the providers and the market. Effectively, one of the information sources is 'redundant', resulting in a welfare loss.

1.5.2 Two-Period Problem with Two Information Sources

Lastly, we turn to the model with two periods and two information sources, in order to investigate the effects of the intertemporal incentives in a competitive setting.

For viewer i, the expected utility of not acquiring information, acquiring information once from the R-biased source and acquiring information twice from an R-biased source are given by $u^0(i)$, $u^1_R(i,c,\lambda_R)$ and $u^2_R(i,c,\lambda_R)$, respectively, in section 4 and acquiring information once from the L-biased source is given by $u^1_L(i,c,\lambda_L)$ in subsection 5.1. The expected utility of viewer i from acquiring signals twice from the L-biased source is given by

$$\begin{aligned} u^2_L(i,c,\lambda_L) =& [(1-i)+i(1-\lambda_L)^2] \\ & \left(\max\left\{1-\frac{2(1-\lambda_L)^2 i}{1-(2-\lambda_L)\lambda_L i}, \frac{2(1-\lambda_L)^2 i}{1-(2-\lambda_L)\lambda_L i}-1\right\}-2c\right) \\ & + i\lambda_L(1-c)+i(1-\lambda_L)\lambda_L(1-2c) \end{aligned}$$

where the first term is her expected utility when she does not receive a conclusive signal, the second term is her expected utility when she receives a conclusive signal in the first period and the last term is her expected utility when she receives a conclusive signal in the second period (as noted above, when acquiring information from an L-biased information source, the conclusive signal is s_r^L and the non-conclusive signal is s_l^L).

Now we turn to the central question of this subsection: Under what conditions can we sustain a fully informative equilibrium, and how do these conditions differ from the ones we found in subsection 4.2? In order to understand and answer these questions, we provide the following definition.

Definition 1.5.4. *An equilibrium of the game is called a full information equilibrium if and only if either $\lambda_R^* = 1$ or $\lambda_L^* = 1$ or both.*

The definition above outlines what a fully informative equilibrium is in our model. We say that an equilibrium is fully informative if and only if at least one of the information provides perfectly informative signals. In such a case, all viewers have an option to learn the state of the world from that provider's signal and end their learning at the end of period one (if they decide to learn at all). Notice that we can sustain a fully informative equilibrium in three distinct ways: *(a)* both information sources provide fully informative signals, or *(b)* only one of the information sources provide fully informative signals.

The result below pins down the set of all possible fully informative equilibria.

Theorem 1.5.5. $\lambda_R^* = \lambda_L^* = 1$ *constitutes the unique full information equilibrium of the game with two periods and two information sources.*

The result above clearly demonstrates the effects of competition on the intertemporal incentives of the information sources. By section 4.2, we know that, under a monopoly scenario, an information source provides full information under specific values of c and a. Namely, the source provides fully informative signals if and only if either c is relatively large or the value of c is at an intermediate level and a is relatively low. This implies that the source does not provide fully informative signals when c is relatively low or c is at an intermediary level and a is relatively high, both of which are possible under the parametric restrictions we set when we laid out the model. The result above shows that when we add another information source to the model that is possibly biased in the opposite direction, the incentives created by competition completely eliminates these incentives created by multiple periods, and as a result the case where both information sources provide fully informative signals is an equilibrium of the model, for any value of c or a, given our assumptions.

Furthermore, the result shows that both providers sending perfectly informative signals is the unique full information equilibrium. In other words, there cannot be any other equilibrium where at least one of the information providers reveal the state of the world perfectly. This is again a direct result of the competitive forces in the market. Lastly, the result shows uniqueness in the space of full information equilibria. It is important to note that there might still exist equilibria where both information sources provide signals that are not fully informative. Furthermore, notice that the unique full information equilibrium is still wasteful in the sense that there are two information providers who produce perfectly informative signals, but only one is enough to reveal the state of the world.

1.6 Conclusion

In this paper, we provided a series of models that investigate bias in information market. More specifically, we showed that biased information providers can occur in equilibrium even without relying on external incentives or behavioral anomalies. In the models, we showed that, the intertemporal incentives of the information providers can lead to such a result. The information providers, when interacting with viewers who are facing an optimal stopping problem, can strategically introduce bias into their signals in order to retain some portion of the viewers in the future and earn a higher payoff. Furthermore, we showed that competition can reverse this result and lead to a fully informative equilibrium. However, importantly, the unique fully informative equilibrium in such a case is wasteful since all the providers send fully informative signals, incurring the cost.

The aim of this paper was to investigate the effects of intertemporal incentives and competition on the information providers' biases under the simplest possible formulation. There are many avenues of extension for future work. First and foremost, it would certainly be an important contribution to pin down the full set of equilibria of the model presented in subsection 5.2, the problem with two periods and two information sources. Notice that the result we present only says that there is a unique fully informative equilibrium, but we do not present any results regarding the equilibria where no signal is fully informative which might exist. It is our conjecture, though, that such equilibria do not exist due to supermodular nature of the game, but the specifics require a more fully-formed model than what we provide in this paper.

Another interesting extension would be letting the static parameters vary with decision variables. For example, by allowing different prior distributions for the viewers, one can investigate the effects of societal polarization on these results. Furthermore, the attention cost can also be an increasing function of the informativeness of the received signal, which can be interpreted as signals that are more informative are usually more costly to understand since they might be more convoluted compared to the less informative ones. These changes would interact with the intertemporal incentives of the information provider and might result in some important comparative statics results.

Furthermore, we introduce competition in a stripped down manner, just by adding an opposite-biased information provider. There almost certainly is value in trying to understand the effects of further competition, such as understanding how many information providers can participate in a market and how that number changes with respect to the parameters of the model. It is our conjecture that the unique full information equilibrium still exists regardless of the number of information providers, albeit with a maximum amount of providers depending on the parameters.

Another important aspect is the cost function of the information providers, which again is not our main objective to understand in this paper. We make some simple assumptions regarding the form and the parameters of the function and move to the results. But it is clear from the results and proofs provided in the appendix that a lot can possibly change with a different form for the cost function. Hence, allowing for a more abstract cost function and understanding how the results depend on it can uncover important effects on the results, especially considering a heterogeneous information market where some information providers have a different cost function than the others.

All in all, the paper provides a model with rational information providers and viewers but biased signals in equilibrium. The literature on information markets historically relied on external sources of incentives or behavioral biases to explain the empirical observation of the existence of biased information sources in real life. This paper provides an explanation as to why that could be the case without relying on such assumptions. There could be more explanations for this empirical phenomenon that are in line with neoclassical economic theory, and the literature could benefit from the articulation of these possible explanations.

Appendix

Proof of Lemma 1.4.1. (i) First assume $\lambda_R \geq c$. Notice that $\frac{2(1-\lambda_R)+c}{2(2-\lambda_R)}$ is continuous and decreasing in λ_R and $1-\frac{c}{2\lambda_R}$ is continuous and increasing in λ_R. Since the lower bound of $\mathcal{A}^R$ is decreasing and the upper bound is increasing in λ_R. Since λ_R has no effect on $\mathcal{A}^R$ or creates a jump from an empty set value to a non-empty set value for $\mathcal{A}^R$ when $\lambda_R < c$, this concludes the proof.

(ii) Notice that $\frac{2(1-\lambda_R)+c}{2(2-\lambda_R)}$ is continuous and increasing in c and $1-\frac{c}{2\lambda_R}$ is continuous and decreasing in c. Since the lower bound of $\mathcal{A}^R$ is increasing and the upper bound is decreasing in c, these mean that $|\mathcal{A}^R|$ is decreasing in c. Furthermore, an increase in c contracts the domain of non-empty attention set. These observations together conclude the proof. ■

Proof of Theorem 1.4.2. The result directly follows from the fact that $1-\frac{c}{2\lambda_R}-\frac{2(1-\lambda_R)+c}{2(2-\lambda_R)}-a\lambda_R^2$ is strictly increasing in λ_R when $\lambda_R \geq c$, $c < \frac{1}{2}$ and $a < \frac{1-c}{4}$. Notice that the derivative of π^R with respect to λ_R is given by

$$\frac{\partial \pi^R}{\partial \lambda_R} = \begin{cases} \frac{1}{2}\left(\frac{2-c}{(2-\lambda_R)^2}\frac{c}{\lambda_R^2} - 4a\lambda_R\right), & \text{if } c \leq \lambda_R \leq 1 \\ -2a\lambda_R, & \text{otherwise} \end{cases}$$

Notice that $\frac{\partial \pi^R}{\partial \lambda_R} > 0$ when $\lambda_R \geq c$ given our assumption $a < \frac{1-c}{4}$. This leads to $\lambda_R^* = 1$. ■

Proof of Lemma 1.4.3. (i) Notice that $|\mathcal{A}_1^R|$ is equal to $|\mathcal{A}^R|$ when $c < \widehat{c}$ and $c \leq \lambda_R < \underline{\lambda}_R$, or $c < \widehat{c}$ and $\overline{\lambda}_R \leq \lambda_R \leq 1$, or $c \geq \widehat{c}$ and $c \leq \lambda_R \leq 1$. Hence, as pinned down in the proof of Lemma 1.4.1, it is decreasing in c under these parameter values. Furthermore, notice that $1-\frac{2(1-c)}{4-(4+c-2\lambda_R)\lambda_R}$ is increasing in c as well. Furthermore, it is greater than $\frac{2(1-\lambda_R)+c}{2(2-\lambda_R)}$ when $c = \widehat{c}$ for any value of λ_R. This concludes the proof.

(ii) Again, notice that $|\mathcal{A}_1^R|$ is equal to $|\mathcal{A}^R|$ when $c < \widehat{c}$ and $c \leq \lambda_R < \underline{\lambda}_R$, or $c < \widehat{c}$ and $\overline{\lambda}_R \leq \lambda_R \leq 1$, or $c \geq \widehat{c}$ and $c \leq \lambda_R \leq 1$. Hence, as pinned down

in the proof of Lemma 1.4.1, it is increasing in λ_R under these parameter values. Furthermore, $1 - \frac{2(1-c)}{4-(4+c-2\lambda_R)\lambda_R}$ is decreasing in λ_R as well. Furthermore, they are equal to each other when $\lambda_R = \underline{\lambda}_R$ and $\lambda_R = \overline{\lambda}R$, which concludes the proof.

(iii) We established that $1 - \frac{2(1-c)}{4-(4+c-2\lambda_R)\lambda_R}$ is increasing in c in part *(i)*. Furthermore, notice that $\frac{(2\lambda_R-c)(1-\lambda_R)}{(2-2\lambda_R+c)\lambda_R}$ is decreasing in c. This concludes the proof.

(iv) We established that $1 - \frac{2(1-c)}{4-(4+c-2\lambda_R)\lambda_R}$ is decreasing in λ_R in part *(ii)*. Now we turn to the upper bound. Notice that $\frac{(2\lambda_R-c)(1-\lambda_R)}{(2-2\lambda_R+c)\lambda_R}$ is increasing in λ_R if and only if $\lambda_R \leq \frac{c+2}{4}$, since its derivative with respect to λ_R is $\frac{c(c-4\lambda_R+2)}{(c-2\lambda_R+2)^2\lambda_R^2}$. This means that $|\mathcal{A}_2^R|$ is increasing in λ_R when $\lambda_R \leq \frac{c+2}{4}$. The derivative of $|\mathcal{A}_2^R|$ with respect to λ_R is $-\frac{4c}{(c+2)(c-2\lambda_R+2)^2} + \frac{c}{(c+2)\lambda_R^2} + \frac{2(1-c)(c+4-4\lambda_R)}{(4-(c+4)\lambda_R+2\lambda_R^2)^2}$, which is monotonically decreasing in λ_R and is equal to 0 when λ_R is equal to the root of the following polynomial (which is unique in $\lambda_R \in (c,1)$) : $-16(2-c)\lambda_R^5 + (96+16c-20c^2)\lambda_R^4 + (-96-144c+8c^2+8c^3)\lambda_R^3 + (32+200c+56c^2-4c^3-c^4)\lambda_R^2 + (-128c-48c^2-8c^3)\lambda_R + 32c + 16c^2 = 0$. Labelling this root $\widehat{\lambda}_R$ concludes the proof. ■

Proof of Theorem 1.4.4. First, notice that $1 - \frac{c}{2\lambda_R} - \frac{2(1-\lambda_R)+c}{2(2-\lambda_R)} - 2a\lambda_R^2$ is strictly increasing in λ_R. Furthermore, $\left(2 - \frac{\lambda_R}{2}\right)\frac{2(1-c)}{4-(4+c-2\lambda_R)\lambda_R} - \frac{c}{2\lambda_R} + \left(1 - \frac{\lambda_R}{2}\right)\left(\frac{(2\lambda_R-c)(1-\lambda_R)}{(2-2\lambda_R+c)\lambda_R} - 1\right) - 2a\lambda_R^2$ is strictly concave in λ_R. Denote the maximizing argument of this function as λ_R'. Notice that as $c \to 0$, $\underline{\lambda}_R \to 0$ and $\overline{\lambda}_R \to 1$. This means that we are in the second case given in π^R above, and $\lambda_R^* = \lambda_R'$ for $c \to 0$. Then, as we increase c, the first case becomes feasible as well. Furthermore, notice that the maximum value function is strictly decreasing in c, but this decrease is faster for the concave function than the strictly increasing function since the derivative of the first with respect to c is lower than the derivative of the second with respect to c. This means that there exists a $\underline{c}_1$ such that $\lambda_R^* < 1$ for any $c < \underline{c}_1$, where $\underline{c}_1$ is pinned down by

$$\left(2 - \frac{\lambda_R'}{2}\right)\frac{2(1-\underline{c}_1)}{4-(4+\underline{c}_1-2\lambda_R')\lambda_R'} - \frac{\underline{c}_1}{2\lambda_R'} + \left(1 - \frac{\lambda_R'}{2}\right)\left(\frac{(2\lambda_R'-\underline{c}_1)(1-\lambda_R')}{(2-2\lambda_R'+\underline{c}_1)\lambda_R'} - 1\right) = 1 - \underline{c}_1$$

which is the payoff equality condition for $\lambda_R^* = \lambda_R'$ and $\lambda_R^* = 1$ at $a = 0$. For any $c < \underline{c}_1$, the payoff to the information provider is always higher when $\lambda_R = \lambda_R'$ compared to $\lambda_R = 1$, regardless of the value of a. For $c > \underline{c}_1$, λ_R^* is equal to either λ_R' or 1, depending on the value of a.

Furthermore, since both functions are monotonically decreasing in a, for any given value of c, there exists $\widehat{a}_1$ such that for $a < \widehat{a}_1$ we have $\lambda_R^* = \lambda_R'$ and $\lambda_R^* = 1$ otherwise, where $\widehat{a}_1$ is pinned down by

$$\left(2 - \frac{\lambda_R'}{2}\right) \frac{2(1-c)}{4-(4+c-2\lambda_R')\lambda_R'} - \frac{c}{2\lambda_R'} + \left(1 - \frac{\lambda_R'}{2}\right)\left(\frac{(2\lambda_R' - c)(1-\lambda_R')}{(2-2\lambda_R'+c)\lambda_R'} - 1\right) - 2\widehat{a}_1\lambda_R'^2 = 1 - c - 2\widehat{a}_1$$

Similarly, there exists $\bar{c}_1$ such that $\lambda_R^* = 1$ for any $c > \bar{c}_1$ regardless of the value of a, where $\bar{c}_1$ is pinned down by

$$\left(2 - \frac{\lambda_R'}{2}\right) \frac{2(1-\bar{c}_1)}{4-(4+\bar{c}_1-2\lambda_R')\lambda_R'} - \frac{\bar{c}_1}{2\lambda_R'} + \left(1 - \frac{\lambda_R'}{2}\right)\left(\frac{(2\lambda_R' - \bar{c}_1)(1-\lambda_R')}{(2-2\lambda_R'+\bar{c}_1)\lambda_R'} - 1\right) - \frac{1-\bar{c}_1}{4}\lambda_R'^2 = \frac{1-\bar{c}_1}{4}$$

which is the payoff equality condition for $\lambda_R^* = \lambda_R'$ and $\lambda_R^* = 1$ at $a = \frac{1-c}{4}$, the highest value of a.

Both are less than $\widehat{c}$, since otherwise the payoff equality condition would not hold, which concludes the proof. ■

Proof of Lemma 1.5.1. *(i)* Notice that $1 - \frac{c}{2\lambda_R}$ is strictly increasing in λ_R and both $\frac{2(1-\lambda_R)+c}{2(2-\lambda_R)}$ and $\frac{1-\lambda_R}{2-\lambda_R-\lambda_L}$ are strictly decreasing in λ_R, which yields the desired result.

(ii) Notice that $\frac{1-\lambda_R}{2-\lambda_R-\lambda_L}$ is increasing in λ_L, which yields the desired result.

(iii) Notice that $1 - \frac{c}{2\lambda_R}$ is strictly decreasing in c and $\frac{2(1-\lambda_R)+c}{2(2-\lambda_R)}$ is strictly increasing in c, which yields the desired result.

(iv) Notice that $\frac{c}{2\lambda_L}$ is strictly decreasing in λ_L and both $\frac{2-c}{2(2-\lambda_L)}$ and $\frac{1-\lambda_R}{2-\lambda_R-\lambda_L}$ are decreasing in λ_L, which yields the desired result.

(v) Notice that $\frac{1-\lambda_R}{2-\lambda_R-\lambda_L}$ is decreasing in λ_R, which yields the desired result.

(vi) Notice that $\frac{c}{2\lambda_L}$ is strictly increasing in c and $\frac{2-c}{2(2-\lambda_L)}$ is strictly decreasing in c, which yields the desired result. ■

Proof of Theorem 1.5.2. Notice that both π^R and π^L are continuous, and strictly increasing except when they are equal to $-a\lambda^2$, in which case $\lambda = 0$ is optimal. Furthermore, notice that both π^R and π^L attain strictly positive values when $\lambda_R = \lambda_L = 1$, which concludes the proof. ■

Proof of Lemma 1.5.3. Notice that under the equilibria of the both models, the viewers with priors in $\left[\frac{c}{2}, 1 - \frac{c}{2}\right]$ acquire signals that perfectly reveal the state of the world and the the viewers with priors in $\left[0, \frac{c}{2}\right]$ and $\left[1 - \frac{c}{2}, 1\right]$ do not acquire a signal. Hence there is no change in the total viewer welfare in the equilibria of the two models. In the model with one source, the R-biased information source receives a payoff of $1-c-a$ which is also equal to the total provider payoff. In the model with two sources, both providers receive payoff of $\frac{1-c}{2}-a$, which implies that the total provider payoff is $1-c-2a$. Since the total provider payoff decreases by a as a result of competition and the total viewer payoff stays the same, we reach to the conclusion in the statement of the lemma. ■

Proof of Theorem 1.5.5. First, we show that $\lambda_R^* = \lambda_L^* = 1$ constitutes an equilibrium of the game. When the signals are fully informative from both information sources, the viewer with initial priors $\left[\frac{c}{2}, 1 - \frac{c}{s}\right]$ acquire information once, and are indifferent between the sources, hence the sources equally share the viewers. In such a case, the profit of an information source is given by $\frac{1-c}{2} - 2a$. If an information source deviates from this action profile, i.e. provides less than fully informative signals, no viewer acquires information from that source, hence its profit becomes $-2a\lambda$, which is maximized at $\lambda = 0$. Hence, $\lambda_R^* = \lambda_L^* = 1$ can be sustained in equilibrium as long as $\frac{1-c}{2} - 2a \geq 0$, which implies $a \leq \frac{1-c}{4}$, which is true by assumption.

Now consider an equilibrium where $\lambda_R^* = 1$ and $\lambda_L^* \in (0, 1)$. In such a case, the L-biased information source's payoff is $-2a\lambda_L$, which is strictly negative. Hence there exists a profitable deviation for the L-biased information source, which is $\lambda_L = 0$. Similarly, we cannot sustain an equilibrium where $\lambda_L^* = 1$ and $\lambda_R^* \in (0, 1)$.

Hence, $\lambda_R^* = \lambda_L^* = 1$ is an equilibrium of the game (and, by definition, is fully informative), and no other full information action profile can be sustained as a part of the perfect Bayesian equilibrium of the game. ■

Chapter 2

Voter Behavior and Information Aggregation in Elections with Supermajority

2.1 Introduction

Supermajority rules are prevalent in electoral democracies around the world even though their benefits are sometimes questioned. The main argument in favor of such rules is the assumption that they foster compromise among different political factions. But in some cases the voters do not have a say in these compromise outcomes, and cannot directly vote for them. In this paper, we develop a model of elections with heterogeneous voters to analyze the effects of such supermajority rules on voter behavior. Our main objective is to understand whether such institutions achieve what they are trying to achieve, as well as to investigate the voter incentives under these rules. More specifically, we examine the effects of various parameters, such as the level of the supermajority threshold or the value of the compromise outcome, on the voter behavior and the probability of the compromise outcome being implemented. We also investigate the information aggregation properties of supermajority rules. This is an important question since elections are perceived as effective only if they can aggregate the information that is dispersed among the voters.

Supermajority requirements are prevalent around the world and have a history that dates back to the Roman Empire (Schwartzberg, 2013). For example, most of the countries around the world have rules in place for constitutional amendments that require some form a supermajority rule (Schwartzberg, 2010), a type of so-called entrenchment clauses (Hein, 2020). A prominent and somewhat infamous example of supermajority institutions is the ones that exist in the United States governance systems, on both federal (Kim, 1993) and state[1] levels. As one of the most prominent political institutions in the United States, the Congress is curbed with various supermajoritarian institutions at many levels of its conduct as well. Arguably the most famous one of these rules is Rule XXII of the Senate[2], also known as the cloture rule, which requires a three-fifths majority to end discussion on a bill or motion and move on to the voting. If such a supermajority cannot be sustained, then the bill 'dies' in the waiting list due to the Senate's tradition of unlimited debate. Effectively, this means that a minority of 41 senators can keep most pieces of legislation on the Senate floor for an 'endless debate'. This is more commonly known as the filibuster.

The filibuster rule has been the center of many political debates in the United States, especially in the last twenty years. Politicians and activists who are on the progressive end of the political spectrum voiced negative opinions about the rule, since many bills that can create meaningful change are filibustered. The most prominent argument for maintaining Rule XXII is that the slow-moving political cogs of the United States government were intentional, instituted by the United States constitution. Such a system is said to engage different views, resulting in a compromise[3]. Kyrsten Sinema, a Senator from Arizona, said "The

[1]https://www.latimes.com/la-oe-goldberg22-2009mar22-story.html, https://www.ncsl.org/research/fiscal-policy/supermajority-vote-requirements-to-pass-the-budget635542510.aspx

[2]The full text of Rule XXII of the Senate can be found at https://www.rules.senate.gov/rules-of-the-senate

[3]https://douthat.blogs.nytimes.com/2010/01/04/bushs-supermajorities/,

idea of the filibuster was created by those who came before us in the United States Senate to create comity and to encourage senators to find bipartisanship and work together."[4] Hence, the justification of Rule XXII is that a non-conclusive vote tally, or the possibility of it, leads to legislative propositions that ponder to both sides of the aisle. More often than not, the compromise of a bill is already in the works when it is proposed, if the expectation is that it will fail to create a three-fifths majority. A recent example of such a process is the minimum wage debate that has been taking place since 2019, which aims to increase the federal minimum wage to $15 an hour[5]. Even though there has not been any vote on a bill, there is already an alternative compromise legislation, proposing to increase the minimum wage to $10 or $11 an hour. Another example is the recent gun control bill, which was introduced as a compromise because a stronger version was known to be impossible to pass[6]. It is unclear how the existence of such compromises, a natural by-product of the supermajority institutions, affect the incentives of the senators when they are voting on a legislation. Moreover, as noted by legal scholars (Chemerinsky & Fisk, 1997), the compromises are usually the outcomes of opaque processes that take place in the Senate Committees. Hence, given that the compromises are exogenous to most of the voters in the Senate, it is not obvious that a higher supermajority threshold leads to a higher probability of a compromise outcome.

Supermajority institutions exist in the governance structures of various multiparty democracies as well. For example, when there are multiple parties in a contest to form a government, the absolute majority rule becomes a supermajority rule. In many countries with a multiparty political system, there is an absolute majority rule to form a government, which can be an almost impossible threshold to clear. In these systems, some of the parties must band together and compro-

https://www.nytimes.com/2019/03/05/opinion/oppression-majority.html

[4]https://www.washingtonpost.com/outlook/2021/06/15/filibuster-bipartisanship-manchin-sinema/

[5]https://www.thenation.com/article/economy/federal-minimum-wage-stimulus/

[6]https://www.bbc.com/news/world-us-canada-61919752

mise, form what is called a coalition government, to secure a confidence vote that requires an absolute majority. In some counties, such as Germany, coalition governments are the norm. More importantly, the voters in such elections cannot vote for a specific coalition government, whose policy proposals can differ tremendously from the individual parties that constitute them. They are the outcome of either pre- or post-elections processes. In effect, there are outcomes of the elections that the voters cannot explicitly vote for.

In order to formalize these ideas, in the model below, we consider a population of voters who have two options to vote for, but there are three distinct outcomes for the election. More specifically, the voters cannot explicitly vote for the compromise outcome; it can occur only if neither of the alternatives reaches to the supermajority threshold. This is an important component of our model that separates it from the preceding studies on supermajority elections.

Equipped with this formulation of a supermajority rule, we consider a voting game that takes place in a two-group, two-state, two-option but three-outcome world. The voters, who belong to one of the two preference groups, can vote for either of the two options, but the implemented outcome might be the compromise, which is not on the ballot. To set the ideas, consider a world with states L and R, groups of voters with state dependent heterogeneous preferences τ_L and τ_R, but three possible outcomes, with slight abuse of notation, L, C and R. The voters can explicitly vote only for L and R. Voters in τ_L prefer L when the state is L, but their preferences differ when the state is R. Similarly, voters in τ_R prefer R when the state is R, but their preferences differ when the state is L. The supermajority threshold is q, which means that L or R is implemented if and only if at least q share of the voters vote for them. If neither L nor R gets q share of the votes, then the outcome is C. Notice that, in such an environment, there are two cases where a voter is potentially pivotal, the case where her vote is pivotal for the decision between L and C, which we call $L-\textit{pivotality}$, and the case where her vote is pivotal for the decision between R and C, which we call $R-\textit{pivotality}$.

The symmetricity that comes from the fact that the supermajority threshold is the same for both L and R leads to most of the results being dependent on whether the expected vote share of an outcome is more than half or not. This holds because voters face symmetric pivotality events, and the expected vote share makes one of them more likely than the other. In turn, voters behave as if the more likely pivotality event is going to happen. It is important to note that these results have testable policy implications, because the expected vote share is observable in the real world.

We first take a voting game where there is a relatively small electorate who know the state of the world, and analyze the effects of the crucial components of the model, such as the supermajority threshold and the compromise value, on the individual voter behavior and the outcome of the election. We then move on to the setting where the number of voters grow to infinity, and observe if and how the results change. We then make the state unknown, and carry out the same initial investigation. Lastly, we again incorporate infinitely many voters to understand the information aggregation properties of elections with supermajority rule.

One of the most important results of the paper is that, in line with the popular argument, a higher supermajority threshold leads to a higher probability of a compromise outcome. Even though the outcome is not surprising, we offer a deeper understanding of the effects of the supermajority threshold. More specifically, we identify two channels through which supermajority threshold changes the probability of an outcome: the direct effect through the needed level of votes, and an indirect one through the equilibrium effects on the expected vote share. We show that both of these effects align and lead to an increase in the probability of a compromise outcome.

As another crucial result, we look at the information aggregation properties of the supermajoritarian elections as defined by the paper, and show that there is no information aggregation. More specifically, given any election with supermajority rule, we can find preference distributions for both groups in the electorate such that the outcome of the election under incomplete information does not converge to the outcome of the election under complete information as the number of voters grow large. This is an important result because, as argued by many in the information aggregation literature, one of the crucial roles of elections in our society is to aggregate important information that is dispersed in the population, and they are reliable to lead to welfare maximizing outcomes to the extent that they correctly aggregate information.

The structure of the paper is as follows. We first provide the model preliminaries. In subsection 2, we provide a literature review on the topics related to this paper, such as supermajority elections, information aggregation and studies that incorporate multiple pivotal events in elections. We then provide a stripped down version of the model to illustrate the effects of multiple pivotal events on voter behavior. Then we move on to the first formulation, which is elections (i.e. finite number of voters) under perfect information, presented in subsection 5. In this subsection, we investigate the properties of supermajority elections with small (i.e. finite) and large (i.e. infinite) electorates. We then move on to elections under incomplete information in subsection 5, where we investigate the information aggregation properties of supermajoritarian elections. The paper concludes with subsection 7. All the proofs are relegated to the appendix.

2.2 Related Literature

This paper contributes to several strands of literature on elections: (i) information aggregation in elections, (ii) supermajority rules and (iii) elections with multiple pivotal events. There are also a few papers that pin down specific settings where information aggregation fails.

The research on the information aggregation properties of elections goes back to the seminal work by Condorcet (1785), now known as the Condorcet Jury Theorem, which states that as long as each voter is more likely to vote for the 'right outcome' (in the societal welfare maximizing sense) and the voters have common goals, then as the electorate gets larger, the right outcome will prevail in a majoritarian election, hence large electorates are less likely to make a wrong decision than a single decision maker. But Condorcet assumes that each voter decides on her vote 'sincerely'; they base their action solely on their own information. Austen-Smith & Banks (1996) showed that sincere voting is usually irrational; it does not constitute a Nash equilibrium of the voting game. Wit (1998) derives the important equilibria of the voting game, as studied by Austen-Smith & Banks (1996), and shows that the Condorcet Jury Theorem holds when we allow for mixed voting strategies. Then, a plethora of papers extended the results of the Condorcet Jury Theorem to more general assumptions regarding information structures, voter preferences, voting rules. Feddersen & Pesendorfer (1997) studied information aggregation in a strategic setting and showed that information aggregation still occurs in two-candidate elections with noisy private information, under both simple majority and supermajority rules. This setting is obviously different than ours, since the model presented in this paper incorporates more than two candidates, and the voter cannot explicitly vote for one. Myerson (1998a) extended this result to Poisson elections where the number of voters is uncertain. Duggan & Martinelli (2001) further extended this result to the settings with a continuum of signals. This paper is an addition to this literature that provides a negative result; a specific setting where information aggregation is violated.

A subgroup within this line of research on strategic voting and incomplete information features models where the voters have to consider multiple pivotal events instead of one. This occurs when either *(i)* there are either more than two outcomes, or *(ii)* there are separate electoral requirements for distinct groups of voters or *(iii)* there are multiple rounds of voting. For example, Myatt (2007) also considers three possible outcomes where the voters can only vote for two of them. But the reason why is that the non-votable outcome in their setting is interpreted as status quo and is uniformly disliked by every voter. In our model, some voters prefer the non-votable outcome, which we interpret as compromise. They also do not comment on information aggregation properties of such elections. In another important paper, Ekmekci (2009) considers a plurality voting model with three candidates with a potentially biased endorser. In their formulation, the endorser has information regarding the probability distribution of the preferences of the voters, which are otherwise unknown. Hence, endorsements have the potential to be coordination signals for the voters, even when the bias of the endorser is known by the voters. Clearly, their paper is significantly different than ours, since the voting rules are different, and ours does not incorporate an endorser. Another similar paper is Myatt (2017) where they consider three possible outcomes for an election where voters can either vote for a candidate or cast a protest vote, and the possible outcomes are candidate loses, candidate wins and protest fail, and candidate wins and protest succeeds, where whether the protest succeeds depends on the share of the protest votes. Their setup is different than ours primarily because they assume uniform preferences among voters where everybody prefers the outcome where the candidate wins and the protest succeeds, along with other differences. Moreover, they do not investigate the information aggregation properties of such elections. Alternatively, Maug & Yilmaz (2002) consider an environment where there are two distinct classes of voters who vote on an issue, and the result depends on the vote share surpassing a predetermined level in both classes (similar to bicameral voting). They investigate the information aggregation properties in this environment and find that two-class voting systems outperform (in the social welfare maximizing sense) single class systems as long as there is conflict of interest between the two classes of voter groups. Again this setting is different than ours because it does not incorporate an outcome that the voters cannot explicitly vote for. Bouton & Castanheira (2012) consider an environment with three possible outcomes, where the majority of the electorate is divided between two alternatives and there is a minority of the voters who back a third alternative. This third alternative is strictly inferior for the majority. Then they look at the tradeoff of aggregating information and coordinating to defeat the inferior candidate for the majority of the electorate. Bouton (2013) looks at the run-off electoral systems where the candidates with the two highest vote share move on to the second round of votes. Another similar paper is Bouton & Gratton (2015). This is clearly different than our single-round election model. They uncover several results about run-off elections, but they do not consider information aggregation.

Our paper differs from most of the literature on supermajority rules because there is no status quo bias. When supermajority rules are incorporated into the election models, it is usually the case that there is an asymmetric threshold: An alternative must pass the supermajority vote share to get implemented, otherwise the status quo remains. Hence they have an inherent 'conservative bias' (Goodin & List, 2006). Our model contains no status quo option. Even if we interpret the compromise outcome as the status quo (it prevails if none of the non-compromise outcomes secure the required supermajority of the votes), the alternatives do not have an asymmetric threshold to pass to get implemented in this formulation of supermajority rules.

There are several specific settings of elections where the Condorcet Jury Theorem does not hold, as pinned down by the literature. For example, Mandler (2012) shows that when the precision of the signals is uncertain, information aggregation fails. Bhattacharya (2013) pins down a condition called 'strong preference monotonicity' and shows that if preference distributions violate this rule, then information aggregations fails for any voting rule. The setting of this result allows for only two alternatives.

2.3 A Motivating Example

In this section, we provide a straightforward and stripped-down example of the problem at hand to motivate the theoretical investigation we are conducting in the rest of the paper.

Suppose an agent is one of the 11 people on the newly-formed board of directors of a company and has to vote on the pay increase for the CEO. Their options are either no increase at all or doubling the compensation of the chief executive. Since the members assumed their roles recently, they do not know much about each other, except that there are two possible types of members: friendly with the CEO and neutral towards the CEO.

A board member's utility from the outcome is measured by the distance between her preferred CEO pay change and the realized outcome, which is either %0 or %200. She and other directors know that the friendly members always unanimously vote for doubling the compensation if the stock price increased in the previous year, and some of them still might vote for the same option when the stock price drops, depending on their preferences of a percentage pay increase, known to be normally distributed with mean %75 and standard deviation %5[7]. On the other hand, the neutral members always unanimously vote for no increase in the compensation if the stock price fell last year, and some of them still might vote for the same option when the stock price goes up, depending on their preferences of a percentage pay increase, known to be normally distributed with mean %25 and standard deviation %5. It is also commonly known that there is a 0.5 probability that any given member of the board is friendly with the CEO. The decision of the board is reached if one of the two options gain at least 9 votes. If none reaches to that number, then it is assumed by all of the members that there will be a default increase in line with the inflation expectations, which is %50. Notice that this is one of the fundamental aspects that sets this paper apart from much of the existing literature on elections with supermajority rules.

Consider the problem of a board member who is neutral towards the CEP. Given that the stock price went up last year and her ideal CEO compensation increase is %110, how should she vote? Notice that if she could vote directly for %50 increase she would, but she cannot, since only available options for voting are

[7]For the sake of the argument, allow for negative changes in the compensation, i.e., wage cuts.

%0 and %200. If she were to pick her action disregarding the details of the voting rule (also called voting 'sincerely' (Austen-Smith & Banks, 1996)), she would vote for doubling the CEO compensation. But a rational voter considers the pivotal events in the election and votes accordingly. There are two pivotal events for this agent: The event where her and eight other board members vote for no increase (and 2 others vote for doubling the compensation) and the event where her and eight other board members vote for doubling the compensation (and 2 others vote for no increase). Only in these two events she casts a vote that is outcome relevant. Notice that this is another one of the fundamental aspects that sets this paper apart along with the feature mentioned above.

As it usually is the case in the voting literature, her voting strategy is going to be defined by a cutoff parameter: If her preferred pay increase of %110 is above that cutoff value, she is going to vote for doubling the pay, and if it is below the cutoff value, she is going to vote for no pay increase. Moreover, there is a cutoff value where an agent is going to be indifferent between voting for an increase and no increase in the CEO pay. Call that cutoff value c.

Given this cutoff value, the probability of a randomly selected board member voting for doubling the pay can be calculated as follows: With probability one half the member is friendly with the CEO and hence votes for a pay increase for sure, and with probability one half the member is neutral and votes for an increase only if her preference parameter is lower than the cutoff value, pinned down by the probability $F(c|0.25, 0.05^2) = \frac{1}{2}[1 + \text{erf}(\frac{c-0.25}{0.05\sqrt{2}})]$. Hence the probability of a randomly selected board member voting for a pay increase, which is equal to the expected pay increase vote share, is $\frac{1}{2} + \frac{1}{4}[1 + \text{erf}(\frac{c-0.25}{0.05\sqrt{2}})]$. Denote this value by i_c. Given this value, the probability that a voter is pivotal for the pay increase outcome is $\binom{10}{8} i_c^8 (1 - i_c)^2$. Similarly, the probability that a voter is pivotal for the no pay increase outcome is $\binom{10}{2} i_c^2 (1 - i_c)^8$.

By definition, an agent with the cutoff value is going to be indifferent between the two options. This implies that her expected utility from voting for an increase, given by

$$\binom{10}{8} i_c^8 (1-i_c)^2(-|2-c|) + \binom{10}{2} i_c^2 (1-i_c)^8(-|0.5-c|)$$

Similarly, her expected utility from voting for no increase is given by

$$\binom{10}{8} i_c^8 (1-i_c)^2(-|0.5-c|) + \binom{10}{2} i_c^2 (1-i_c)^8(-|c|)$$

Given these, we can back out the cutoff value that is pinned down by the indifference condition. Equating these two expressions and solving for c yields $c = 1.25$. This implies that if an agent's preferred CEO pay increase is less than %125, she is going to vote for no increase, and if an agent's preferred CEO pay increase is more than %125, she is going to vote for doubling the pay. Given this cutoff value, it is straightforward that our agent is going to vote for no increase in the CEO pay, which is different than her vote if she were to vote sincerely.

2.4 The Model Preliminaries

There are $n+1$ voters denoted by i, who can take the action $a_i \in \{l, r\}$, which we call voting l or voting r. There are two states of the world: $\omega \in \{U, D\} = \Omega$. As mentioned before, one of the leading differences of this paper from the existing literature is that there is an outcome that the voters cannot explicitly vote for. More specifically, the election outcome is $j \in \{L, C, R\} \in \mathbb{R}$ and we assume that $L < C < R$. In line with the observations in Section 1, the outcome that voters cannot explicitly vote for is the outcome C, which we label as the compromise outcome because by assumption its value is in between the outcomes L and R.

The outcome of the election is determined by the following procedure: For a given threshold $q > \frac{1}{2}$, if the number of l votes is greater than qn then the outcome is L, if it is between qn and $(1-q)n$ then the outcome is C and if it is fewer than

$(1-q)n$, then the outcome is R [8]. Importantly, the voters cannot explicitly vote for the outcome C, it occurs only if the vote tallies fall on a specific interval, determined by the parameter q. We call this procedure a *supermajority rule.*

There are two preference groups $\tau \in \{\tau_L, \tau_R\} = \mathcal{T}$, with $P(\tau_L) = \rho$. Prior to the elections, each agent knows her group, and has a *preference type*, $x \sim \mathcal{N}(\mu_x, \sigma_x^2)$ for τ_L and $y \sim \mathcal{N}(\mu_y, \sigma_y^2)$ for τ_R. We label x's probability distribution function and cumulative distribution function as f_x and F_x, respectively, and y's probability distribution function and cumulative distribution function as f_y and F_y, respectively. We interpret $y_U - \mu_X$ and $\mu_y - x_D$ as the measures of disagreement across groups in states U and D respectively, and the group variances as the measure of disagreement within that group. The preference parameter of an agent is state dependent. For the agents in the group τ_L, if the state is U, the preference parameter x_U is equal to the preference type x, and if the state is D, the preference parameter is $x_D < L$. For the agents in the group τ_R, if the state is D, the preference parameter y_D is equal to the preference type y, and if the state is U, the preference parameter is $y_U > R$. Hence τ_L type agents have heterogeneous preferences only in state U whereas τ_R type agents have heterogeneous preferences only in state D.

A voter's utility from casting a vote depends on her type, her preference parameter (which depends on the state) and the outcome, and is given by

$$u_L(j, \omega) = -(x_\omega - j)^2$$

$$u_R(j, \omega) = -(y_\omega - j)^2$$

for τ_L and τ_R voters, respectively. As a result, when the state is D, all τ_L voters prefer the outcome L and when the state is U, all τ_R voters prefer the outcome R. Hence, agents have single peaked preferences around their state dependent bliss points.

[8]For simplicity, we consider only the populations such that qn (and hence, $(1-q)n$) is an integer.

We mostly concern ourselves with τ_L type voters in state U when we are investigating the properties of the equilibrium under known state, since the corresponding results for the state D and τ_R type voters are similar.

2.5 Elections with Known State

In this section, we assume that the state is announced before the election takes place and present the existence of the equilibrium, and then investigate its various properties of interest.

A voting strategy for voter i, $\pi_i : \Omega \times \mathcal{T} \times \mathbb{R} \to [0, 1]$, is a measurable function from the state, her group and preference type to the probability of voting l. Let $\overline{\pi}$ be the corresponding strategy profile for all voters. An equilibrium is defined as a symmetric Nash equilibrium of the voting game specified above. Given a symmetric strategy profile $\overline{\pi}$, define the probability that a randomly selected voter votes l in state ω as

$$t(\omega, \overline{\pi}) = \rho \int_{-\infty}^{\infty} \overline{\pi}(\omega, x; \tau_L) f_x(x) dx + (1 - \rho) \int_{-\infty}^{\infty} \overline{\pi}(\omega, y; \tau_R) f_y(y) dy \tag{2.1}$$

There are two cases where a voter can influence the outcome of the election: the case where exactly qn voters vote l, where her vote will lead to either L or C, and the case where exactly $(1 - q)n$ voters vote l, where her vote will lead to either R or C. We refer to these events as $L - pivotality$ and $R - pivotality$ and denote them by piv_L and piv_R, respectively. Given a symmetric strategy profile $\overline{\pi}$ and threshold q, the probability that a vote is L-pivotal in state ω is

$$P(piv_L | \omega, \overline{\pi}) = \binom{n}{qn} t(\omega, \overline{\pi})^{qn} (1 - t(\omega, \overline{\pi}))^{(1-q)n}$$

and the probability that a vote is R-pivotal in state ω is

$$P(piv_R|\omega,\overline{\pi}) = \binom{n}{(1-q)n} t(\omega,\overline{\pi})^{(1-q)n}(1-t(\omega,\overline{\pi}))^{qn}$$

We mostly omit $\overline{\pi}$ from the notation unless it is needed for clarity.

Simple algebra shows that when the state is D, all τ_L voters prefer to vote l. If the state is U, then a τ_L voter i votes l if and only if

$$\frac{P(piv_R|U)}{P(piv_L|U)} = \left(\frac{1-t(U)}{t(U)}\right)^{(2q-1)n} \geq (\leq) \frac{(2x_{U,i}-L-C)(C-L)}{(C+R-2x_{U,i})(R-C)} \tag{2.2}$$

when $x_U < (>) \frac{R+C}{2}$. Similarly, if the state is U, all τ_R voters prefer to vote r. If the state is D, then they vote r if and only if

$$\frac{P(piv_L|D)}{P(piv_R|D)} = \left(\frac{t(D)}{1-t(D)}\right)^{(2q-1)n} \geq (\leq) \frac{(R+C-2y_{D,i})(R-C)}{(2y_{D,i}-L-C)(C-L)} \tag{2.3}$$

when $y_D > (<) \frac{C+L}{2}$. When the state is known, the voters simply compare the ratio of the probability of the two pivotality events to the ratio of the payoffs in those respective events, and decide accordingly.

The following result establishes the existence of the equilibrium of the known-state voting game. The equilibrium strategies have a cutoff structure where the voters vote l if their preference parameter is below some value, and vote r otherwise, if the state is the one in which they have a preference heterogeneity.

Lemma 2.5.1. *There exist cutoff types $\widehat{x}_U, \widehat{y}_D \in (\frac{C+L}{2}, \frac{R+C}{2})$ such that the strategy profiles*

$$\pi_R(\omega, y_{D,i}) = \begin{cases} 0 & \textit{if } \omega = U \textit{ or } \omega = D \textit{ and } y_{D,i} > \widehat{y}_D \\ 1 & \textit{if } \omega = D \textit{ and } y_{D,i} < \widehat{y}_D \end{cases}$$

$$\pi_L(\omega, x_{U,i}) = \begin{cases} 0 & \textit{if } \omega = U \textit{ and } x_{U,i} > \widehat{x}_U \\ 1 & \textit{if } \omega = D \textit{ or } \omega = U \textit{ and } x_{U,i} < \widehat{x}_U \end{cases}$$

constitute an equilibrium for the known-state voting game. The resulting expected vote share is

$$t(\omega, \overline{\pi}) = \begin{cases} \rho F_x(\widehat{x}_U) & \text{if } \omega = U \\ \rho + (1-\rho) F_y(\widehat{y}_D) & \text{if } \omega = D \end{cases}$$

The lemma above, while pinning down the existence of the cutoff equilibrium, also informs us about the nature of the cutoff voter preference. Notice that, since each group's cutoff voters' preferences belong in $(\frac{C+L}{2}, \frac{R+C}{2})$, the voters with the cutoff preference necessarily strictly prefers C over R or L, regardless of her group. This is expected since any voter with preference parameter that is less than $\frac{C+L}{2}$ strictly prefers L to any other outcome, and any voter with preference parameter that is more than $\frac{R+C}{2}$ strictly prefers R to any other outcome, hence they all have dominant actions l and r, respectively. If the voter with the cutoff preference parameter could vote directly for C she would, but she cannot. Hence, she wants to vote l in the case where she is R-pivotal and r in the case where she is L-pivotal to tip the scale towards the compromise outcome. Anticipating the resulting vote shares, she finds which pivotal event is more likely, and votes accordingly.

Notice that there is a close relationship between the location of $\widehat{x}_U$ and the equilibrium expected vote share. First, as expected, the equilibrium expected vote shares are increasing in cutoff points of the preference groups. More specifically, the cutoff preference parameter for τ_L type voters is going to be closer to R than L if the expected l vote share is less than half, regardless of the value of C, and vice versa. The value of C, on the other hand, determines how close the cutoff parameter is going to be to one of its border values R and L. In addition to C, q and n affect this distance as well.

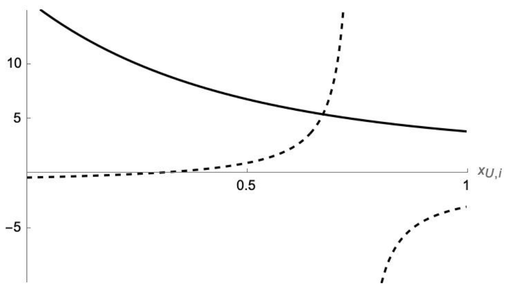

Figure 2.1: The plot of the left hand side (solid) and the right hand side (dashed) of equation 2 when $\rho = 0.4$, $\mu_x = 0$, $\sigma_x = 1$, $q = 0.6$, $n = 10$, $R = 1$, $C = 0.5$ and $L = 0$. The intersection point pins down $\widehat{x}_U$.

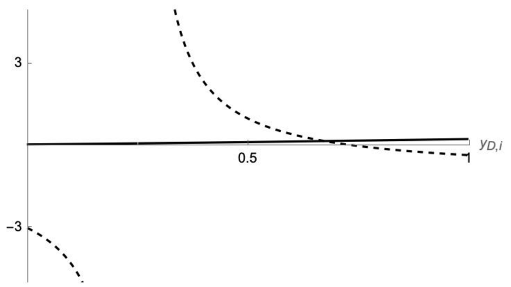

Figure 2.2: The plot of the left hand side (solid) and the right hand side (dashed) of equation 3 when $\rho = 0.4$, $\mu_y = 0$, $\sigma_y = 1$, $q = 0.6$, $n = 10$, $R = 1$, $C = 0.5$ and $L = 0$. The intersection point pins down $\widehat{y}_D$.

In order to demonstrate this, let $\widehat{\Gamma}(t(U), q, n) \equiv \left(\frac{1-t(U)}{t(U)}\right)^{(2q-1)n}$. From inequality 2.2, this implies that

$$\widehat{x}_U = \frac{\widehat{\Gamma}R^2 + (1-\widehat{\Gamma})C^2 - L^2}{2(\widehat{\Gamma}R + (1-\widehat{\Gamma})C - L)}$$

Since $\widehat{\Gamma} > 1$ when $t(U) < \frac{1}{2}$, this implies that $\widehat{x}_U > \frac{R+L}{2}$, and similarly when $t(U) > \frac{1}{2}$, this implies that $\widehat{x}_U < \frac{R+L}{2}$.

Given the existence of these cutoff types, we turn to investigating how the cutoff types change with respect to the parameters of the model. The lemma below establishes some of the important properties of the cutoff preference parameter for agents in group τ_L with respect to some important parameters of the model. Note that most of these result depend on the level of the expected vote share, which is an empirically observable quantity.

Lemma 2.5.2. *$\widehat{x}_U$ is*

(i) increasing in μ_x,

(ii) increasing in σ_x if and only if $t(U) > \frac{\rho}{2}$,

(iii) increasing in q if and only if $t(U) < \frac{1}{2}$,

(iv) increasing in n if and only if $t(U) < \frac{1}{2}$.

(v) *increasing in C if and only if $\widehat{x}_U > C$ when $t(U) < \frac{1}{2}$ and $\widehat{x}_U < C$ when $t(U) > \frac{1}{2}$.*

Some of the results are somewhat unexpected. For example, the cutoff preference parameter for a τ_L voter increases as q increases when the expected l vote share is less than half. To see why, consider the incentives of the cutoff voter. She is indifferent by definition, she would vote r if she was L-pivotal and l if she was R-pivotal. When $t(U) < \frac{1}{2}$, as q increases, the relative probability of being R-pivotal as opposed to L-pivotal increases. Hence, since R-pivotality event is seen as more likely after the threshold increase, she is more incentivized to vote l to tip the outcome towards C in the $R-$pivotality event. This result is somewhat unexpected because it implies that an increase in the threshold level can lead to a lower probability of a randomly selected voter voting l when $t(U) > \frac{1}{2}$. This is due to the fact that people condition their strategies on the relative likelihood of two pivotal events.

Similarly, the cutoff preference parameter for a τ_L voter increases as C gets closer to R, when the expected l vote share is less than half and the cutoff is higher than C. A change in C affects the cutoff type through two channels; one is the change in the payoffs from the outcomes and the other is the change in the expected l vote share. Again, consider the incentives of the cutoff voter. She is indifferent by definition, she would vote r if she was L-pivotal and l if she was R-pivotal. An increase in C increases her payoff from outcome C since $\widehat{x}_U > C$. Hence, she is more incentivized to vote l, because she prefers the outcome C even more in the R-pivotal event, which is more likely to occur since $t(U) < \frac{1}{2}$. Notice is that the same is not true when $\widehat{x}_U < C$, since in such a case an increase in C leads to a decrease in her payoff. Hence the overall effect depends on both the proximity of the preference parameter to R and L, and also whether it is below or above C. This result also seems counterintuitive because it implies that if the compromise value gets closer to R, the probability that a randomly selected voter voting l might decrease. This is also due to the fact that people condition their strategies on the relative likelihood of two pivotal events.

One of the most important questions that we are concerned with is the effects of model parameters on the probability of the outcomes. Since the action taken, i.e. the vote of each agent, is a binary variable, the probability of an outcome follows a binomial distribution where the parameters are the number of voters and the probability of voting l. Hence, given the equilibrium cutoffs, we can define the probability of the outcome being R as

$$\Pi_R(\omega) = \begin{cases} \sum_{k=0}^{(1-q)n-1} \binom{n}{k} (\rho + (1-\rho)F_y(\widehat{y}_D))^k (1-\rho-(1-\rho)F_y(\widehat{y}_D))^{n-k} & \text{if } \omega = D \\ \sum_{k=0}^{(1-q)n-1} \binom{n}{k} (\rho F_x(\widehat{x}_U))^k (1-\rho F_x(\widehat{x}_U))^{n-k} & \text{if } \omega = U \end{cases}$$

and similarly the probability of the outcome being L as

$$\Pi_L(\omega) = \begin{cases} \sum_{k=qn+1}^{n} \binom{n}{k} (\rho + (1-\rho)F_y(\widehat{y}_D))^k (1-\rho-(1-\rho)F_y(\widehat{y}_D))^{n-k} & \text{if } \omega = D \\ \sum_{k=qn+1}^{n} \binom{n}{k} (\rho F_x(\widehat{x}_U))^k (1-\rho F_x(\widehat{x}_U))^{n-k} & \text{if } \omega = U \end{cases}$$

As a result, we have the probability of the outcome being C as

$$\Pi_C(\omega) = \begin{cases} \sum_{k=(1-q)n}^{qn} \binom{n}{k} (\rho + (1-\rho)F_y(\widehat{y}_D))^k (1-\rho-(1-\rho)F_y(\widehat{y}_D))^{n-k} & \text{if } \omega = D \\ \sum_{k=(1-q)n}^{qn} \binom{n}{k} (\rho F_x(\widehat{x}_U))^k (1-\rho F_x(\widehat{x}_U))^{n-k} & \text{if } \omega = U \end{cases}$$

Given these probabilities, we investigate how the probability of a compromise outcome changes with respect to some important parameters of the model. The following auxiliary result will be useful for our understanding of the question. Any change in any of the parameters of the voting game has an effect on the equilibrium expected vote share, which in turn affects the outcome probabilities.

Lemma 2.5.3. *(i)* $\Pi_C(U)$ *is increasing in* $t(U)$ *if and only if* $t(U) < \frac{1}{2}$.

(ii) $\Pi_L(U)$ *is increasing in* $t(U)$.

(ii) $\Pi_R(U)$ *is decreasing in* $t(U)$.

First, we look at the effect of within-group disagreement on the probability of the outcome C being implemented, presented in the lemma below.

Lemma 2.5.4. Π_C *is increasing in* σ_x *if and only if* $t(U) > \frac{1}{2}$.

We find that the probability of a compromise outcome increases with respect to within-group disagreement for the τ_L group if and only if the expected l vote share is more than half. Remember that the cutoff preference parameter is increasing in within-group disagreement if the expected l vote share is more than $\frac{\varrho}{2}$, which means more than half of τ_L group voters are voting l since no τ_R group voter is voting l. But also, an increase in σ_x can be interpreted as the density of τ_L type voters over from the middle of the preference spectrum decreasing, and correspondingly the density of τ_L type voters over the edges of the preference spectrum increasing. The expected vote share then governs if these replaced voters are still voting the same or if they are switching to to the other vote. If the expected l vote share is more than $\frac{\varrho}{2}$, then some of the τ_L type voters in the middle of the preference spectrum moves towards R, meaning the cutoff parameter is decreasing. Since fewer people are voting l, the probability of the outcome L decreases, and correspondingly the probability of outcome C increases. Interestingly, which of these two effects dominates depends on the expected vote share being more or less than $\frac{1}{2}$. But since the effect through the second channel is minimal when $t(U)$ is around $\frac{\varrho}{2}$, the first effect overtakes the second one for $t(U) \in (\frac{\varrho}{2}, \frac{1}{2})$.

We now turn to the effect of across group disagreement, measured by the distance between μ_x and y_U when the state is U, on the probability of outcome C, presented in the lemma below.

Lemma 2.5.5. Π_C *is increasing in* μ_x *if and only if* $t(U) > \frac{1}{2}$.

Notice that an increase in μ_x implies a decrease in across-group disagreement. Hence we can conclude that an increase in the across-group disagreement leads to an increase in the probability of outcome C being implemented when $t(U) < \frac{1}{2}$, and vice versa. To see why, notice that an increase in across-group disagreement in state U happens when the average τ_L voter moves even further to the left, which has a negative effect on the probability of the outcome C, but this leads to a leftward shift in the cutoff preference parameter, which leads to an increase in the probability of the outcome C. Which one of these two contradicting effects prevails depends on the expected l vote share. If it is more than half, the first one prevails, leading to an overall decrease in the probability of outcome C as a result of a decrease in μ_x, and if it is less than half, the second one prevails, leading to an overall increase in the probability of outcome C as a result of a decrease in μ_x.

As we mentioned in the introductory section of the paper, one of the most prominent arguments in favor of supermajority rules is that it leads to compromise outcomes. In order to check if that is indeed the case, we now investigate the effects of a change in the threshold value on the probability of C being implemented.

Theorem 2.5.6. Π_C *is increasing in* q*.*

The result confirms the widespread argument that supermajority rules promote compromise outcomes. Even though the result above is in line with expectations, it is not a straightforward result. A change in q affects the equilibrium of the voting game through two channels. First, we have what can be called the direct channel: Since q is increasing, it is harder to reach the supermajority threshold. This intuitively increases the probability of a compromise outcome. Second, we have the indirect channel: A change in q changes the expected vote share and the location of the cutoff preference parameter, hence the equilibrium strategy in the voting game. More precisely, as we demonstrated in Lemma 2.5.2, an increase in q moves the cutoff voter parameter to the left if and only if the expected l vote share is less than half. First, assume that the expected l vote share is less than half. This leads to an increase in the cutoff preference parameter, leading to an increase in the expected vote share. Since the expected vote share is assumed to be less than half, this means the probability of compromise increases. Second, assume that the expected vote share is more than half. This leads to a decrease in the cutoff preference parameter, leading to a decrease in the expected vote share.

Since the expected vote share is assumed to be more than half, this means the probability of compromise increases again. As a result, the secondary effect of an increase in q through the change in expected vote share is always positive. When combined, these two channels imply that an increase in q leads to an increase in the probability of a compromise outcome. Hence, we are able to confirm the common assumption that supermajority rules lead to compromise outcomes, even though the reasoning is more intricate than expected.

We now turn to the question of how changes in the compromise value affect the compromise probability, presented in the lemma below.

Lemma 2.5.7. Π_C *is increasing in* C *if and only if* $\widehat{x}_U > C$.

The result turns out to be quite straightforward. If the compromise value moves closer to R, when the compromise probability increases if and only if the cutoff voter is already to the right of the compromise value. Notice that such a change in C is going to entice the τ_L type voters who are closer to the right of the preference parameter spectrum and who would otherwise vote r, to vote l, and has no effect on the τ_L type voters who are closer to the left of the preference spectrum, since they have no other choice than voting l.

2.5.1 Electorate Welfare in Small Elections with Known State

In this subsection, we investigate the effects of changes in the supermajority threshold q and the compromise value C on the welfare of τ_L type voters in elections under state U. This is important because just because an increase in threshold parameter leads to an increase in the probability of the compromise outcome does not mean that the voters are going to benefit from such a change.

The expected utility of a τ_L type voter with a generic preference parameter x_U

from the elections under state U is

$$\mathbb{E}[u_L(j,U)] = -\Big[\Pi_C(U,\widehat{x}_U)(x_U - C)^2 + \Pi_R(U,\widehat{x}_U)(x_U - R)^2 + \Pi_L(U,\widehat{x}_U)(x_U - L)^2\Big]$$

First, we consider the individual welfare, and investigate if the effects are qualitatively the same on the voters. In order to answer those questions, we first need to identify some properties of Π_R and Π_L, similar to the results in the previous subsection.

Lemma 2.5.8. *(i)* $\Pi_R(U,\widehat{x}_U)$ *is decreasing in* q *if* $t(U) < \frac{1}{2}$.

(ii)$\Pi_L(U,\widehat{x}_U)$ *is decreasing in* q *if* $t(U) > \frac{1}{2}$.

Equipped with this auxiliary result, we first investigate the effects of an increase in the supermajority threshold q on the welfare of the voters. The results are presented in the lemma below.

Lemma 2.5.9. *The welfare of the* τ_L *voters who vote* l *decreases as a result of an increase in* q *when* $t(U) < \frac{1}{2}$.

The results says that the welfare of the τ_L voters who would otherwise vote l decreases as a result of an increase in the supermajority threshold when the expected l vote is less than half. Intuitively, if the expected l vote is less than half, then an increase in the supermajority threshold decreases the probability of the outcome C and the outcome L at the same time, leading to a welfare loss. When the expected l vote is more than half, the outcome is ambiguous because an increase in the supermajority threshold increases the probability of the outcome C but decreases the probability of outcome L.

2.5.2 Large Elections with Known State

In this subsection, we investigate the equilibrium of the voting game as the number of voters grows without bound. The limiting game is important because it will serve as the comparison benchmark for the information aggregation result. In line with the existing literature, we say that there is information aggregation if the equilibrium of the incomplete information voting game implements the same outcome as the game where the state is announced before the elections, as the number of voters goes to infinity. This idea will be formalized in the relevant subsection later. Hence, it is elemental to pin down the equilibrium of the known state voting game as n goes to infinity.

The following result establishes the equilibrium of the limit voting game as n goes to infinity.

Lemma 2.5.10. *Let $\widehat{x}_{U,n}$ and $\widehat{y}_{D,n}$ be the cutoff types that constitute an equilibrium of the known-state voting game with n voters. As $n \to \infty$*

$$\widehat{x}_{U,n} \to \begin{cases} \frac{C+L}{2} & \text{if } F_x(\frac{C+L}{2}) \geq \frac{1}{2\rho} \\ F_x^{-1}(\frac{1}{2\rho}) & \text{if } F_x(\frac{C+L}{2}) < \frac{1}{2\rho} < F_x(\frac{R+C}{2}) \\ \frac{R+C}{2} & \text{if } F_x(\frac{R+C}{2}) \leq \frac{1}{2\rho} \end{cases}$$

and

$$\widehat{y}_{D,n} \to \begin{cases} \frac{C+L}{2} & \text{if } F_y(\frac{C+L}{2}) > \frac{\frac{1}{2}-\rho}{1-\rho} \\ F_y^{-1}(\frac{\frac{1}{2}-\rho}{1-\rho}) & \text{if } F_y(\frac{C+L}{2}) < \frac{\frac{1}{2}-\rho}{1-\rho} < F_x(\frac{R+C}{2}) \\ \frac{R+C}{2} & \text{if } F_y(\frac{R+C}{2}) < \frac{\frac{1}{2}-\rho}{1-\rho} \end{cases}$$

The resulting vote share is

$$t(\omega, \overline{\pi}) = \begin{cases} \rho F_x(\frac{C+L}{2}) & \text{if } F_x(\frac{C+L}{2}) > \frac{1}{2\rho} \text{ and } \omega = U \\ \frac{1}{2} & \text{if } F_x(\frac{C+L}{2}) < \frac{1}{2\rho} < F_x(\frac{R+C}{2}) \text{ and } \omega = U \\ \rho F_x(\frac{R+C}{2}) & \text{if } F_x(\frac{R+C}{2}) < \frac{1}{2\rho} \text{ and } \omega = U \\ \rho + (1-\rho) F_y(\frac{C+L}{2}) & \text{if } F_y(\frac{C+L}{2}) > \frac{\frac{1}{2}-\rho}{1-\rho} \text{ and } \omega = D \\ \frac{1}{2} & \text{if } F_y(\frac{C+L}{2}) < \frac{\frac{1}{2}-\rho}{1-\rho} < F_x(\frac{R+C}{2}) \text{ and } \omega = D \text{ and} \\ \rho + (1-\rho) F_y(\frac{R+C}{2}) & \text{if } F_y(\frac{R+C}{2}) < \frac{\frac{1}{2}-\rho}{1-\rho} \text{ and } \omega = D \end{cases}$$

Notice that this result enables us to understand the preference location of the cutoff voters and demonstrates how they try to lead to equal vote shares $t(U) = t(D) = \frac{1}{2}$. Consider group τ_L (the case for group τ_R is similar). If the measure of the τ_L voters who strictly prefer outcome L to any other outcome is enough to form a simple majority, then the cutoff preference parameter converges to $\frac{C+L}{2}$, which means that everyone who is indifference or strict preference towards outcomes C or R votes r, to make sure that the outcome is not L. Similarly, if the measure of the τ_L voters who strictly prefer outcome R to any other outcome is enough to form a simple majority, then the cutoff preference parameter converges to $\frac{R+C}{2}$, which means that everyone who is indifference or strict preference towards outcomes L or C votes l, to make sure that the outcome is not R. If neither is the case and the voters with strict preference towards L or R are not enough to form a simple majority, then the cutoff voter preference parameter converges to a value that makes sure that the outcome of the vote is equal, i.e. both options get half the votes.

Given this result, we can uncover when specific outcomes are implemented as a function of the model parameters, which are presented in the lemma below.

Lemma 2.5.11. *When the state is U, R is implemented if and only if $F_x(\frac{R+C}{2}) < \frac{1-q}{\rho}$. Otherwise, C is always implemented. When the state is D, L is implemented if and only if $F_y(\frac{C+L}{2}) > \frac{q-\rho}{1-\rho}$. Otherwise, C is always implemented.*

Notice that the result above outlines the effects of preference polarization in large elections. First, take a voting game where the electorate preferences are *aligned* in the sense that in a given state, the preferences of both τ_L and τ_R are same on average, with little variance, i.e., $\mu_x = y_U$, $\mu_y = x_D$ and σ_x, σ_y are small. The result above implies that, with an aligned electorate, the outcome is C regardless of the state. For comparison, take an electorate with preferences that are *polarized* in the sense that the preferences of each group do not depend on the state on average, with little variance, i.e., $\mu_x = x_D$, $\mu_y = y_U$ and σ_x, σ_y are small. Similarly, the result above implies that, with such an electorate, the outcome of the game is R when the state is U and L when the state is D. Hence, even with supermajoritarian institutions, polarized preferences lead to polarized outcomes.

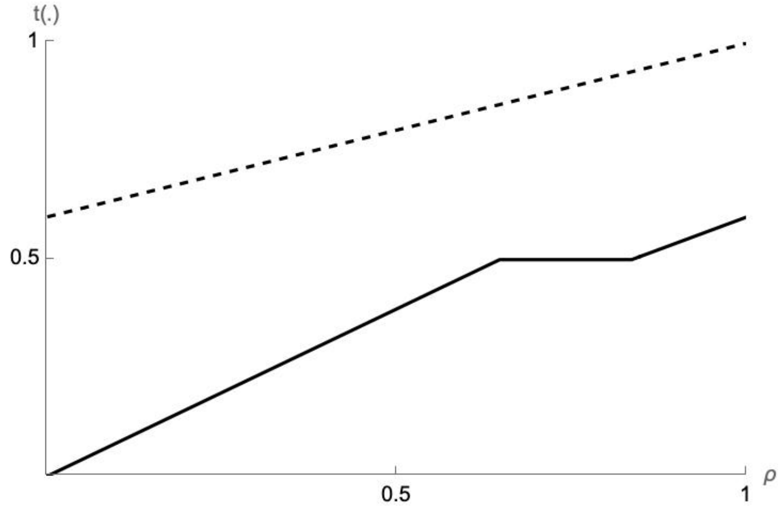

Figure 2.3: The figure plots $t(U)$ (solid) and $t(D)$ (dashed) as a function of ρ when $\mu_x = \mu_y = 0$, $\sigma_x = \sigma_y = 1$, $L = 0$, $C = 0.5$ and $R = 1$.

2.6 Elections with Unknown State

In this section, we analyze the related voting game where the state is not revealed to the voters before the elections. In line with the many existing papers on the topic, we assume that each agent receives a signal about the state of the world, then she forms a posterior belief according to the Bayes' rule. Our main objective is to investigate whether there is information aggregation or not. But the first order of business is to establish the existence of the equilibrium for the incomplete information voting game.

Each agent receives a private signal $s_i \in \{u, d\} = S$, distributed i.i.d. with $P(u|U) = \alpha > \frac{1}{2}$ and $P(d|D) = \beta > \frac{1}{2}$. A voting strategy for voter i, $\pi_i : S \times \mathcal{T} \times \mathbb{R} \to [0, 1]$, is a measurable function from her signal, her group and preference type to the probability of voting l. Let $\overline{\pi}$ be the corresponding strategy profile for all voters. An equilibrium is defined as a symmetric Nash equilibrium of the voting game specified above. Given a symmetric strategy profile $\overline{\pi}$, define the probability that a randomly selected voter votes l in state ω as

$$t(\omega,\overline{\pi}) = \sum_S P(s|\omega)\left(\rho\int_{-\infty}^{\infty}\overline{\pi}(s,x;\tau_L)f_x(x)dx + (1-\rho)\int_{-\infty}^{\infty}\overline{\pi}(s,y;\tau_R)f_y(y)dy\right)$$

As a reminder, given the above $t(\omega,\overline{\pi})$, the probability that a vote is L-pivotal in state ω is

$$P(piv_L|\omega,\overline{\pi}) = \binom{n}{qn}t(\omega,\overline{\pi})^{qn}(1-t(\omega,\overline{\pi}))^{(1-q)n}$$

and the probability that a vote is R-pivotal in state ω is

$$P(piv_R|\omega,\overline{\pi}) = \binom{n}{(1-q)n}t(\omega,\overline{\pi})^{(1-q)n}(1-t(\omega,\overline{\pi}))^{qn}$$

After receiving signal s, an agent's posterior distribution over states conditional on being L-pivotal and receiving signal s is

$$\beta(\omega|piv_L,s) = \frac{P(piv_L|\omega)P(s|\omega)P(\omega)}{\sum_w P(piv_L|w)P(s|w)P(w)}$$

and the posterior belief conditional on being R-pivotal and receiving signal s is

$$\beta(\omega|piv_R,s) = \frac{P(piv_R|\omega)P(s|\omega)P(\omega)}{\sum_w P(piv_R|w)P(s|w)P(w)}$$

After receiving signal s, a τ_L voter votes l if and only if

$$\frac{P(piv_R|D)}{P(piv_L|D)}\frac{R-C}{C-L}\frac{R+C-2x_D}{C+L-2x_D} + \frac{\beta(U|s)}{\beta(D|s)}\frac{P(piv_L|U)}{P(piv_L|D)}\frac{C+L-2x_{U,i}}{C+L-2x_D} + \frac{\beta(U|s)}{\beta(D|s)}\frac{P(piv_R|U)}{P(piv_L|D)}\frac{R-C}{C-L}\frac{R+C-2x_{U,i}}{C+L-2x_D} \geq -1 \quad (2.4)$$

Notice that all τ_L agents with $x_{U,i} < \frac{C+L}{2}$ vote l regardless of their signal.

Similarly, after receiving signal s, a τ_R voter votes r if and only if

$$\frac{P(piv_L|U)}{P(piv_R|U)}\frac{C-L}{R-C}\frac{2y_U-C-L}{2y_U-R-C}+\frac{\beta(D|s)}{\beta(U|s)}\frac{P(piv_R|D)}{P(piv_R|U)}\frac{2y_{D,i}-R-C}{2y_U-R-C} + \frac{\beta(D|s)}{\beta(U|s)}\frac{P(piv_L|D)}{P(piv_R|U)}\frac{C-L}{R-C}\frac{2y_{D,i}-C-L}{2y_U-R-C} \geq -1 \quad (2.5)$$

Notice that all τ_R agents with $y_{D,i} > \frac{R+C}{2}$ vote r regardless of their signal.

The following lemma establishes the existence of the cutoff equilibrium for the elections with unknown state.

Lemma 2.6.1. *There exist cutoff types* $\widehat{x}_{U,d} > \widehat{x}_{U,u} > \frac{C+L}{2}$ *and* $\widehat{y}_{D,u} < \widehat{y}_{D,d} < \frac{R+C}{2}$ *such that*

$$\pi_R(s, y_{D,i}) = \begin{cases} 0 & \text{if } s = u \text{ and } y_{D,i} > \widehat{y}_{D,u} \text{ or } s = d \text{ and } y_{D,i} > \widehat{y}_{D,d} \\ 1 & \text{otherwise} \end{cases}$$

$$\pi_L(s, x_{U,i}) = \begin{cases} 0 & \text{if } s = u \text{ and } x_{U,i} > \widehat{x}_{U,u} \text{ or } s = d \text{ and } x_{U,i} > \widehat{x}_{U,d} \\ 1 & \text{otherwise} \end{cases}$$

constitute an equilibrium of the unknown-state voting game. The resulting vote share is

$$t(\omega, \overline{\pi}) = \begin{cases} \alpha(\rho F_x(\widehat{x}_{U,u}) + (1-\rho)F_y(\widehat{y}_{D,u})) + (1-\alpha)(\rho F_x(\widehat{x}_{U,d}) + (1-\rho)F_y(\widehat{y}_{D,d})) \\ \hfill \text{if } \omega = U \\ (1-\beta)(\rho F_x(\widehat{x}_{U,u}) + (1-\rho)F_y(\widehat{y}_{D,u})) + \beta(\rho F_x(\widehat{x}_{U,d}) + (1-\rho)F_y(\widehat{y}_{D,d})) \\ \hfill \text{if } \omega = D \end{cases}$$

After establishing the existence of the equilibrium and cutoff strategies, we turn to the information aggregation properties of the game at hand. Information aggregation is an important property because one of the main roles of elections, as pointed out by Feddersen & Pesendorfer (1999), is that they aggregate information that is dispersed in the population. For completeness, information aggregation property is defined below, in line with the literature.

Definition 2.6.2. *A sequence of strategy profiles $\overline{\pi}_n$ aggregates information if, as $n \to \infty$, the outcome of the elections with unknown state under $\overline{\pi}_n$ is the same as the outcome of the elections with known state.*

Next, we present the result concerning the information aggregation properties of elections with supermajority rules.

Theorem 2.6.3. *For some preference parameter distributions F_x and F_y, there is no sequence of strategy profiles $\overline{\pi}_n$ that aggregates information.*

It turns out that there is no information aggregation under the supermajority rule, as specified in the model. This brings the advantages of supermajoritarian elections when there are outcomes where the voters cannot explicitly vote for into question. Formally, the proof shows that, for any given set of parameters that pin down a supermajority election, we can find a pair of preference distributions for both groups of the electorate such that the outcome of the elections under incomplete information as the number of voters goes to infinity does not converge to the outcome of the large elections under complete information. More specifically, the proof demonstrates that it is always possible to find a pair of preference distributions F_x and F_y such that, when the state is known, they lead to outcome R when the state is U and outcome C when the state is D, but when the state is unknown, they always lead to outcome C regardless of the true state of the world. In effect, this says that the outcome of the elections would have been different if the voters knew the state of the world, hence implies that the election does not enable the electorate to reach the first-best outcome. The contradictory preference parameter distributions are the ones where there is a majority of the group of voters over a small segment of the parameter domain that is adjacent to $\frac{C+L}{2}$ or $\frac{R+C}{2}$, which are the boundaries of preference parameters that allow for cross-voting. As the number of voters grow, the preference distribution looks more and more like a Dirac delta distribution function. Hence, if there are enough voters in one group of the electorate who technically could vote for the other option but extremely unlikely to do so, then information aggregation fails. This means that information aggregation is usually possible when the probabilities of the parameters are more evenly distributed across the domain.

2.7 Conclusion

In this paper, we constructed a model where there are three possible outcomes but the heterogeneous electorate can only explicitly vote for two of them. We considered supermajority rules where an option gets implemented if and only if the vote share for that option is above a supermajority threshold. If neither option secures such a supermajority, then the outcome that the voters cannot explicitly vote for is implemented. As explained, we interpret this outcome as the compromise outcome, and these elections are similar to the U.S. Senate votes (when the electorate is small) and multi-party elections with absolute majority rules to form a government (when the electorate is large). As previously observed in the literature, a vote is important if and only if it is pivotal for an outcome. Crucially, the elections specified in this paper lead to two distinct but symmetric pivotal events, since there are three outcomes. We conduct a thorough analysis of such elections.

The first important result of the paper is about the relationship between the compromise outcome and the supermajority threshold. One of the most common arguments in favor of the supermajority rules is that they nudge the voters to compromise. In order to verify that, we looked at the effects of an increase in the supermajority threshold on the probability of the compromise outcome. Our result confirms the widespread argument.

The other important result of the paper is about the information aggregation properties of the supermajority elections, as formulated in this study. As observed initially by Condorcet (Austen-Smith & Banks, 1996), elections are efficient as a tool of determining societal outcomes if they aggregate the dispersed information among the electorate. We found that supermajority elections, as formulated in this paper, violate the information aggregation property.

Some future work is needed to generalize the results of this study. One straightforward route is allowing for more than three outcomes, while still maintaining non-votable compromise outcomes. This setting would be especially relevant in the incomplete information world, and would allow us to understand the information aggregation properties better. Another area where the additional research is needed is the implications of the changes in the model parameters other than the supermajority threshold on voter welfare. Since a welfare analysis is not the main point of this paper, we provide a minimal result on the welfare implications of the supermajority threshold. Further research would better our understanding of the effects of supermajority rules on the voters.

Appendix

Proof of Lemma 2.5.1. Existence of $\widehat{x}_U$: The right hand side of inequality 2.2 is continuous and negative when $x_U < \frac{C+L}{2}$ and when $x_U > \frac{R+C}{2}$. Over $(\frac{C+L}{2}, \frac{R+C}{2})$, the expression is positive, strictly increasing, tends to 0 as $x_U \to \frac{C+L}{2}$ and tends to infinity as $x_U \to \frac{R+C}{2}$. Notice that $t(U)$ is increasing in x_U. Then, this means that the left hand side of inequality 2.2 is positive and strictly decreasing over $(-\infty, \infty)$. By the Intermediate Value Theorem, this means that inequality 2.2 must hold with equality for some $\widehat{x}_U \in (\frac{C+L}{2}, \frac{R+C}{2})$. To see that it constitutes a cutoff equilibrium, simple algebra verifies, given that the inequality holds with equality at $\widehat{x}_U$, it is strictly better off to vote l for any τ_L group agent with $x_U < \widehat{x}_U$ and it is strictly better off to vote r for any τ_L group agent with $x_U > \widehat{x}_U$.

Existence of $\widehat{y}_D$: The right hand side of inequality 2.3 is continuous and negative when $y_D > \frac{R+C}{2}$ and when $y_D < \frac{C+L}{2}$. Over $(\frac{C+L}{2}, \frac{R+C}{2})$, the expression is strictly decreasing to 0 . Notice that $t(D)$ is decreasing in y_D. Then, this means that the left hand side of inequality 2.3 is continuous, positive valued and strictly increasing over $(-\infty, \infty)$. By the Intermediate Value Theorem, this means that inequality 2.3 must hold with equality for some $\widehat{y}_D \in (\frac{C+L}{2}, \frac{R+C}{2})$. To see that it constitutes a cutoff equilibrium, simple algebra verifies, given that the inequality holds with equality at $\widehat{y}_D$, it is strictly better off to vote l for any τ_R group agent with $y_D < \widehat{y}_D$ and it is strictly better off to vote r for any τ_R group agent with $y_D > \widehat{y}_D$.

The derivation of $t(\omega, \overline{\pi})$ is straightforward. ■

Proof of Lemma 2.5.2. (i) Notice that $\frac{\partial F_x(\widehat{x}_U)}{\partial \mu_x} = (\frac{\partial \widehat{x}_U}{\partial \mu_x} - 1)\frac{1}{\sigma_x}\phi(\frac{x_U - \mu_x}{\sigma_x})$ where ϕ is the probability distribution function of the standard normal distribution. At $\widehat{x}_U$, inequality 2.2 holds with equality. Taking the total derivative of the equation with respect to μ_x, we get

$$-\frac{(2q-1)n}{\sigma_x F_x(\widehat{x}_U)(1-\rho F_x(\widehat{x}_U))}\left(\frac{1-\rho F_x(\widehat{x}_U)}{\rho F_x(\widehat{x}_U)}\right)^{(2q-1)n}\phi\Big(\frac{\widehat{x}_U-\mu_x}{\sigma_x}\Big)\Big(\frac{\partial \widehat{x}_U}{\partial \mu_x}-1\Big)$$
$$=\frac{2(C-L)(R-L)}{(R-C)(R+C-2\widehat{x}_U)^2}\frac{\partial \widehat{x}_U}{\partial \mu_x}$$

Rearranging, we get

$$\frac{\partial \widehat{x}_U}{\partial \mu_x}=\frac{\frac{(2q-1)n}{\sigma_x F_x(\widehat{x}_U)(1-\rho F_x(\widehat{x}_U))}\left(\frac{1-\rho F_x(\widehat{x}_U)}{\rho F_x(\widehat{x}_U)}\right)^{(2q-1)n}\phi\Big(\frac{\widehat{x}_U-\mu_x}{\sigma_x}\Big)}{\frac{(2q-1)n}{\sigma_x F_x(\widehat{x}_U)(1-\rho F_x(\widehat{x}_U))}\left(\frac{1-\rho F_x(\widehat{x}_U)}{\rho F_x(\widehat{x}_U)}\right)^{(2q-1)n}\phi\Big(\frac{\widehat{x}_U-\mu_x}{\sigma_x}\Big)+\frac{2(C-L)(R-L)}{(R-C)(R+C-2\widehat{x}_U)^2}}$$

which is always positive.

(ii) Notice that $\frac{\partial F_x(\widehat{x}_U)}{\partial \sigma_x}=\phi(\frac{\widehat{x}_U-\mu_x}{\sigma_x})\frac{1}{\sigma_x}(\frac{\partial \widehat{x}_U}{\partial \sigma_x}-\frac{\widehat{x}_U-\mu_x}{\sigma_x})$. At $\widehat{x}_U$, inequality 2.2 holds with equality. Taking the total derivative of the equation with respect to σ_x, we get

$$-\frac{(2q-1)n}{\sigma_x F_x(\widehat{x}_U)(1-\rho F_x(\widehat{x}_U))}\left(\frac{1-\rho F_x(\widehat{x}_U)}{\rho F_x(\widehat{x}_U)}\right)^{(2q-1)n}\phi\Big(\frac{x_U-\mu_x}{\sigma_x}\Big)\Big(\frac{\partial x_U}{\partial \sigma_x}-\frac{x_U-\mu_x}{\sigma_x}\Big)$$
$$=\frac{2(C-L)(R-L)}{(R-C)(R+C-2\widehat{x}_U)^2}\frac{\partial x_U}{\partial \sigma_x}$$

Rearranging, we have

$$\frac{\partial x_U}{\partial \sigma_x}=\frac{x_U-\mu_x}{\sigma_x}\frac{\frac{(2q-1)n}{\sigma_x F_x(\widehat{x}_U)(1-\rho F_x(\widehat{x}_U))}\left(\frac{(1-\rho F_x(\widehat{x}_U)}{F_x(\widehat{x}_U)}\right)^{(2q-1)n}\phi\Big(\frac{x_U-\mu_x}{\sigma_x}\Big)}{\frac{2(C-L)(R-L)}{(R-C)(R+C-2\widehat{x}_U)^2}+\frac{(2q-1)n}{\sigma_x F_x(\widehat{x}_U)(1-\rho F_x(\widehat{x}_U))}\left(\frac{(1-\rho F_x(\widehat{x}_U)}{F_x(\widehat{x}_U)}\right)^{(2q-1)n}\phi\Big(\frac{x_U-\mu_x}{\sigma_x}\Big)}$$

where the second expression on the right hand side is strictly positive. This means that $\frac{\partial x_U}{\partial \sigma_x}>0$ if and only if $\widehat{x}_U>\mu_x$. Since F_x is an increasing function, $\widehat{x}_U>\mu_x$ holds if and only if $F_x(\widehat{x}_U)>F_x(\mu_x)=\frac{1}{2}$. Multiplying both sides with ρ yields the desired result.

(iii) Inequality 2.2 holds with equality at $\widehat{x}_U$. Taking the total derivative of

the both sides with respect to q, we get

$$- \frac{n(\frac{1-\rho F_x(\widehat{x}_U)}{\rho F_x(\widehat{x}_U)})^{(2q-1)n}}{F_x(\widehat{x}_U)(1-\rho F_x(\widehat{x}_U))}\Big((2q-1) f_x(\widehat{x}_U) \frac{\partial \widehat{x}_U}{\partial q} - 2F_x(\widehat{x}_U)(1-\rho F_x(\widehat{x}_U)) \log(\frac{1-\rho F_x(\widehat{x}_U)}{\rho F_x(\widehat{x}_U)})\Big) = \frac{2(C-L)(R-L)}{(R-C)(R+C-2\widehat{x}_U)^2} \frac{\partial \widehat{x}_U}{\partial q}$$

$$- \frac{(2q-1)n f_x(\widehat{x}_U)(\frac{1-\rho F_x(\widehat{x}_U)}{\rho F_x(\widehat{x}_U)})^{(2q-1)n}}{F_x(\widehat{x}_U)(1-\rho F_x(\widehat{x}_U))} \frac{\partial \widehat{x}_U}{\partial q} + n(\frac{1-\rho F_x(\widehat{x}_U)}{\rho F_x(\widehat{x}_U)})^{(2q-1)n)} 2\log(\frac{1-\rho F_x(\widehat{x}_U)}{\rho F_x(\widehat{x}_U)}) = \frac{2(C-L)(R-L)}{(R-C)(R+C-2\widehat{x}_U)^2} \frac{\partial \widehat{x}_U}{\partial q}$$

Rearranging, we get

$$\frac{\partial \widehat{x}_U}{\partial q} = \frac{(\frac{1-\rho F_x(\widehat{x}_U)}{\rho F_x(\widehat{x}_U)})^{(2q-1)n}}{\frac{2q-1}{2}\frac{f_x(\widehat{x}_U)}{F_x(\widehat{x}_U)(1-\rho F_x(\widehat{x}_U))}(\frac{1-\rho F_x(\widehat{x}_U)}{\rho F_x(\widehat{x}_U)})^{(2q-1)n} + \frac{(C-L)(R-L)}{(R-C)(R+C-2\widehat{x}_U)^2}\frac{1}{n}} \log(\frac{1-\rho F_x(\widehat{x}_U)}{\rho F_x(\widehat{x}_U)})$$

which is positive if and only if $\rho F_x(\widehat{x}_U) = t(U) < \frac{1}{2}$.

(iv) Inequality 2.2 holds with equality at $\widehat{x}_U$. Taking the total derivative of the both sides with respect to n, we get

$$(2q-1)\ln\Big(\frac{1-\rho F_x(\widehat{x}_U)}{\rho F_x(\widehat{x}_U)}\Big)\Big(\frac{1-\rho F_x(\widehat{x}_U)}{\rho F_x(\widehat{x}_U)}\Big)^{(2q-1)n} - \frac{(2q-1)n}{\rho\sigma_x F_x(\widehat{x}_U)^2}\Big(\frac{1-\rho F_x(\widehat{x}_U)}{\rho F_x(\widehat{x}_U)}\Big)^{(2q-1)n-1}\phi(\frac{\widehat{x}_U-\mu_x}{\sigma_x})\frac{\partial \widehat{x}_U}{\partial n} = \frac{2(C-L)(R-L)}{(R-C)(R+C-\widehat{x}_U)^2}\frac{\partial \widehat{x}_U}{\partial n}$$

Rearranging, we have

$$\frac{\partial \widehat{x}_U}{\partial n} = \frac{(2q-1)\ln\Big(\frac{1-\rho F_x(\widehat{x}_U)}{\rho F_x(\widehat{x}_U)}\Big)\Big(\frac{1-\rho F_x(\widehat{x}_U)}{\rho F_x(\widehat{x}_U)}\Big)^{(2q-1)n}}{\frac{2(C-L)(R-L)}{(R-C)(R+C-\widehat{x}_U)^2} + \frac{(2q-1)n}{\rho\sigma_x F_x(\widehat{x}_U)^2}\Big(\frac{1-\rho F_x(\widehat{x}_U)}{\rho F_x(\widehat{x}_U)}\Big)^{(2q-1)n-1}\phi(\frac{\widehat{x}_U-\mu_x}{\sigma_x})}$$

Notice that this expression is positive if and only if $\ln\Big(\frac{1-\rho F_x(\widehat{x}_U)}{\rho F_x(\widehat{x}_U)}\Big) > 0$, which holds if and only if $\rho F_x(\widehat{x}_U) < \frac{1}{2}$, which yields the desired result.

(v) Taking the total derivative of the both sides of equation 2.2 with respect to C at $\widehat{x}_U$, we get

$$-\frac{(2q-1)n}{F_x(\widehat{x}_U)(1-\rho F_x(\widehat{x}_U))}\left(\frac{1-\rho F_x(\widehat{x}_U)}{F_x(\widehat{x}_U)}\right)^{(2q-1)n} f_x(\widehat{x}_U)\frac{\partial \widehat{x}_U}{\partial C}$$
$$= \frac{2(R-L)(C-\widehat{x}_U)(2\widehat{x}_U-R-L)}{(R-C)^2(R+C-2\widehat{x}_U)^2} + \frac{2(C-L)(R-L)}{(R-C)(R+C-\widehat{x}_U)^2}\frac{\partial \widehat{x}_U}{\partial C}$$

Rearranging, we get

$$\frac{\partial \widehat{x}_U}{\partial C} = -\frac{\frac{2(R-L)(C-\widehat{x}_U)(2\widehat{x}_U-R-L)}{(R-C)^2(R+C-2\widehat{x}_U)^2}}{\frac{(2q-1)n}{F_x(\widehat{x}_U)(1-\rho F_x(\widehat{x}_U))}\left(\frac{(1-\rho F_x(\widehat{x}_U)}{F_x(\widehat{x}_U)}\right)^{(2q-1)n} f_x(\widehat{x}_U) + \frac{2(C-L)(R-L)}{(R-C)(R+C-\widehat{x}_U)^2}}$$

which is positive if and only if $(C-\widehat{x}_U)(2\widehat{x}_U-R-L) \leq 0$, which requires either $\widehat{x}_U < \min\{C, \frac{R+L}{2}\}$ or $\widehat{x}_U > \max\{C, \frac{R+L}{2}\}$. Moreover, notice that $\widehat{x}_U > \frac{R+L}{2}$ if and only if $t(U) < \frac{1}{2}$. Hence, if $t(U) < \frac{1}{2}$, then $\frac{\partial \widehat{x}_U}{\partial C} > 0$ if and only if $\widehat{x}_U > C > \frac{R+L}{2}$. ■

Proof of 2.5.3. (i) We derive $\frac{\partial \Pi_C}{\partial t(U)}$ by first deriving $\frac{\partial \Pi_C}{\partial F_x(\widehat{x}_U)}$ then dividing it by ρ.

Consider the probability that the number of l votes, denoted by T, votes being equal to exactly t, $P(T=t|F_x(\widehat{x}_U)) = \binom{n}{t}(\rho F_x(\widehat{x}_U))^t(1-\rho F_x(\widehat{x}_U))^{n-t}$. Taking the derivative, we get

$$\frac{\partial}{\partial F_x(\widehat{x}_U)}P(T=t|F_x(\widehat{x}_U))$$
$$= \binom{n}{t}(\rho F_x(\widehat{x}_U))^t(1-\rho F_x(\widehat{x}_U))^{n-t-1}\Big(\frac{t-\rho F_x(\widehat{x}_U)n}{F_x(\widehat{x}_U)}\Big)$$
$$=\binom{n}{t}(\rho F_x(\widehat{x}_U))^t(1-\rho F_x(\widehat{x}_U))^{n-t}\Big(\frac{t-\rho F_x(\widehat{x}_U)n}{F_x(\widehat{x}_U)(1-\rho F_x(\widehat{x}_U))}\Big)$$
$$=\binom{n}{t}(\rho F_x(\widehat{x}_U))^t(1-\rho F_x(\widehat{x}_U))^{n-t}\Big(\frac{t}{F_x(\widehat{x}_U)} + \frac{\rho t}{1-\rho F_x(\widehat{x}_U)}$$
$$- \frac{\rho n}{1-\rho F_x(\widehat{x}_U)}\Big)$$
$$=\binom{n}{t}(\rho F_x(\widehat{x}_U))^t(1-\rho F_x(\widehat{x}_U))^{n-t}\Big(\frac{t}{F_x(\widehat{x}_U)} - \frac{\rho(n-t)}{1-\rho F_x(\widehat{x}_U)}\Big)$$

Then summing over all the values that do not lead to C, we have

$$\frac{\partial}{\partial F_x(\widehat{x}_U)} P((1-q)n \le T \le qn | F_x(\widehat{x}_U))$$
$$= \sum_{t=(1-q)n}^{qn} \binom{n}{t} (\rho F_x(\widehat{x}_U))^t (1-\rho F_x(\widehat{x}_U))^{n-t}$$
$$\Big(\frac{t}{F_x(\widehat{x}_U)} - \frac{\rho(n-t)}{1-\rho F_x(\widehat{x}_U)}\Big)$$
$$= \sum_{t=(1-q)n}^{qn} \binom{n}{t} (\rho F_x(\widehat{x}_U))^t (1-\rho F_x(\widehat{x}_U))^{n-t} \frac{t}{F_x(\widehat{x}_U)}$$
$$- \rho \binom{n}{t} (n-t)(\rho F_x(\widehat{x}_U))^t (1-\rho F_x(\widehat{x}_U))^{n-t-1}$$

Notice that

$$\binom{n}{t}(n-t) = \frac{n!}{t!\,(n-t-1)!} = \frac{(t+1)n!}{(t+1)!\,(n-(t+1))!} = (t+1)\binom{n}{t+1}$$

hence we have

$$\frac{\partial}{\partial F_x(\widehat{x}_U)} P((1-q)n \le T \le qn | F_x(\widehat{x}_U))$$
$$= \sum_{t=(1-q)n}^{qn} \binom{n}{t} (\rho F_x(\widehat{x}_U))^t (1-\rho F_x(\widehat{x}_U))^{n-t} \frac{t}{F_x(\widehat{x}_U)}$$
$$- \rho(t+1)\binom{n}{t+1} (\rho F_x(\widehat{x}_U))^t (1-\rho F_x(\widehat{x}_U))^{n-t-1}$$
$$= \sum_{t=(1-q)n}^{qn} \rho t \binom{n}{t} (\rho F_x(\widehat{x}_U))^{t-1} (1-\rho F_x(\widehat{x}_U))^{n-t}$$
$$- \rho(t+1)\binom{n}{t+1} (\rho F_x(\widehat{x}_U))^t (1-\rho F_x(\widehat{x}_U))^{n-(t+1)}$$

where the second term is the same as the first term except the summation index is shifted forward by 1. Cancelling out the terms, we are left with

$$\begin{aligned}
&\frac{\partial}{\partial F_x(\widehat{x}_U)} P((1-q)n \leq T \leq qn | F_x(\widehat{x}_U)) \\
&= \rho(1-q)n \binom{n}{(1-q)n} (\rho F_x(\widehat{x}_U))^{(1-q)n-1} (1-\rho F_x(\widehat{x}_U))^{qn} \\
-&\rho(qn+1) \binom{n}{qn+1} (\rho F_x(\widehat{x}_U))^{qn} (1-\rho F_x(\widehat{x}_U))^{(1-q)n-1}
\end{aligned}$$

Moreover we have

$$(qn+1)\binom{n}{qn+1} = (qn+1)\frac{n!}{(qn+1)qn!\,((1-q)n-1)!} = (1-q)n\binom{n}{(1-q)n}$$

Plugging this in, we get

$$\begin{aligned}
&\frac{\partial}{\partial F_x(\widehat{x}_U)} P((1-q)n \leq T \leq qn | F_x(\widehat{x}_U)) \\
&= \rho(1-q)n \binom{n}{(1-q)n} (\rho F_x(\widehat{x}_U))^{(1-q)n-1} (1-\rho F_x(\widehat{x}_U))^{qn} \\
-&\rho(1-q)n \binom{n}{(1-q)n} (\rho F_x(\widehat{x}_U))^{qn} (1-\rho F_x(\widehat{x}_U))^{(1-q)n-1} \\
=&\rho(1-q)n \binom{n}{(1-q)n} (\rho F_x(\widehat{x}_U))^{(1-q)n-1} (1-\rho F_x(\widehat{x}_U))^{(1-q)n-1} \\
&\Big((1-\rho F_x(\widehat{x}_U))^{(2q-1)n+1} - (\rho F_x(\widehat{x}_U))^{(2q-1)n+1}\Big)
\end{aligned}$$

which is equal to $\frac{\partial \Pi_C}{\partial F_x(\widehat{x}_U)}$. Hence we have

$$\begin{aligned}
\frac{\partial \Pi_C}{\partial t(U)} = \frac{1}{\rho}\frac{\partial \Pi_C}{\partial F_x(\widehat{x}_U)} =&(1-q)n \binom{n}{(1-q)n} (\rho F_x(\widehat{x}_U))^{(1-q)n-1} (1-\rho F_x(\widehat{x}_U))^{(1-q)n-1} \\
&\Big((1-\rho F_x(\widehat{x}_U))^{(2q-1)n+1} - (\rho F_x(\widehat{x}_U))^{(2q-1)n+1}\Big)
\end{aligned}$$

Simple algebra shows that $\frac{\partial \Pi_C}{\partial t(U)} > 0$ if and only if $t(U) < \frac{1}{2}$.

(ii) Similarly, for L to be implemented, we have

$$\begin{aligned}
&\frac{\partial}{\partial F_x(\widehat{x}_U)} P(T \geq qn+1 | F_x(\widehat{x}_U)) \\
&= \sum_{t=qn+1}^{n} \rho t \binom{n}{t} (\rho F_x(\widehat{x}_U))^{t-1} (1-\rho F_x(\widehat{x}_U))^{n-t} \\
&- \rho(t+1) \binom{n}{t+1} (\rho F_x(\widehat{x}_U))^{t} (1-\rho F_x(\widehat{x}_U))^{n-(t+1)}
\end{aligned}$$

where, again, the second term is the same as the first term except the summation index is shifted forward by 1. Cancelling out the terms, we are left with

$$\begin{aligned}
&\frac{\partial}{\partial F_x(\widehat{x}_U)} P(T \geq qn+1 | F_x(\widehat{x}_U)) \\
&= \rho(qn+1) \binom{n}{qn+1} (\rho F_x(\widehat{x}_U))^{qn} (1-\rho F_x(\widehat{x}_U))^{(1-q)n-1} \\
&- \rho(n+1) \binom{n}{n+1} (\rho F_x(\widehat{x}_U))^{n} (1-\rho F_x(\widehat{x}_U))^{-1} \\
&= \rho(qn+1) \binom{n}{qn+1} (\rho F_x(\widehat{x}_U))^{qn} (1-\rho F_x(\widehat{x}_U))^{(1-q)n-1}
\end{aligned}$$

which is equal to $\frac{\partial \Pi_L}{\partial F_x(\widehat{x}_U)}$. Hence we have

$$\frac{\partial \Pi_L}{\partial t(U)} = \frac{1}{\rho} \frac{\partial \Pi_L}{\partial F_x(\widehat{x}_U)} = (qn+1) \binom{n}{qn+1} (\rho F_x(\widehat{x}_U))^{qn} (1-\rho F_x(\widehat{x}_U))^{(1-q)n-1}$$

(iii) Similarly, for R to be implemented, we have

$$\begin{aligned}
&\frac{\partial}{\partial F_x(\widehat{x}_U)} P(T \leq (1-q)n-1 | F_x(\widehat{x}_U)) \\
&= \sum_{t=0}^{(1-q)n-1} \rho t \binom{n}{t} (\rho F_x(\widehat{x}_U))^{t-1} (1-\rho F_x(\widehat{x}_U))^{n-t} \\
&- \rho(t+1) \binom{n}{t+1} (\rho F_x(\widehat{x}_U))^{t} (1-\rho F_x(\widehat{x}_U))^{n-(t+1)}
\end{aligned}$$

where the second term is the same as the first term except the summation index is shifted forward by 1. Cancelling out the terms, we are left with

$$\frac{\partial}{\partial F_x(\widehat{x}_U)} P(T \leq (1-q)n - 1 | F_x(\widehat{x}_U)) = -\rho(1-q)n \binom{n}{(1-q)n} (\rho F_x(\widehat{x}_U))^{(1-q)n-1}(1-\rho F_x(\widehat{x}_U))^{qn}$$

which is equal to $\frac{\partial \Pi_R}{\partial F_x(\widehat{x}_U)}$. Hence we have

$$\frac{\partial \Pi_R}{\partial t(U)} = \frac{1}{\rho}\frac{\partial \Pi_R}{\partial F_x(\widehat{x}_U)} = -(1-q)n\binom{n}{(1-q)n}(\rho F_x(\widehat{x}_U))^{(1-q)n-1}(1-\rho F_x(\widehat{x}_U))^{qn}$$

■

Proof of Lemma 2.5.4. Taking the derivative of Π_C with respect to σ_x we get $\frac{\partial \Pi_C(U)}{\partial \sigma_x} = \frac{\partial \Pi_C}{\partial F_x(\widehat{x}_U)}\frac{\partial F_x(\widehat{x}_U)}{\partial \sigma_x}$ where $\frac{\partial F_x(\widehat{x}_U)}{\partial \sigma_x} = \phi(\frac{\widehat{x}_U - \mu_x}{\sigma_x})\frac{1}{\sigma_x}(\frac{\partial \widehat{x}_U}{\partial \sigma_x} - \frac{\widehat{x}_U - \mu_x}{\sigma_x})$ and, from Lemma 2.5.3, we have $\frac{\partial \Pi_C}{\partial F_x(\widehat{x}_U)} = \rho(1-q)n\binom{n}{(1-q)n}(\rho F_x(\widehat{x}_U))^{(1-q)n-1}(1-\rho F_x(\widehat{x}_U))^{(1-q)n-1}\Big((1-\rho F_x(\widehat{x}_U))^{(2q-1)n+1} - (\rho F_x(\widehat{x}_U))^{(2q-1)n+1}\Big)$.

Combining everything together, we have $\frac{\partial \Pi_C(U)}{\partial \sigma_x} = \phi(\frac{\widehat{x}_U - \mu_x}{\sigma_x})\frac{1}{\sigma_x}(\frac{\partial \widehat{x}_U}{\partial \sigma_x} - \frac{\widehat{x}_U - \mu_x}{\sigma_x})\rho(1-q)n\binom{n}{(1-q)n}(\rho F_x(\widehat{x}_U))^{(1-q)n-1}(1-\rho F_x(\widehat{x}_U))^{(1-q)n-1}\Big((1-\rho F_x(\widehat{x}_U))^{(2q-1)n+1} - (\rho F_x(\widehat{x}_U))^{(2q-1)n+1}\Big)$

From Lemma 2.5.2, we also know that

$$\frac{\partial x_U}{\partial \sigma_x} = \frac{x_U - \mu_x}{\sigma_x}\frac{\frac{(2q-1)n}{\sigma_x F_x(\widehat{x}_U)(1-\rho F_x(\widehat{x}_U))}\left(\frac{(1-\rho F_x(\widehat{x}_U))}{F_x(\widehat{x}_U)}\right)^{(2q-1)n}\phi\left(\frac{x_U-\mu_x}{\sigma_x}\right)}{\frac{2(C-L)(R-L)}{(R-C)(R+C-2\widehat{x}_U)^2} + \frac{(2q-1)n}{\sigma_x F_x(\widehat{x}_U)(1-\rho F_x(\widehat{x}_U))}\left(\frac{(1-\rho F_x(\widehat{x}_U))}{F_x(\widehat{x}_U)}\right)^{(2q-1)n}\phi\left(\frac{x_U-\mu_x}{\sigma_x}\right)} < \frac{x_U - \mu_x}{\sigma_x}$$

Hence we can conclude that $\frac{\partial \Pi_C(U)}{\partial \sigma_x} > 0$ if and only if $t(U) > \frac{1}{2}$. ■

Proof of Lemma 2.5.5. Taking the derivative of Π_C with respect to μ_x we get $\frac{\partial \Pi_C}{\partial \mu_x} = \frac{\partial \Pi_C}{\partial F_x(\widehat{x}_U)}\frac{\partial F_x(\widehat{x}_U)}{\partial \mu_x}$ where $\frac{\partial F_x(\widehat{x}_U)}{\partial \mu_x} = \phi(\frac{\widehat{x}_U - \mu_x}{\sigma_x})\frac{1}{\sigma_x}(\frac{\partial \widehat{x}_U}{\partial \mu_x} - 1)$ and, from Lemma 2.5.3, we have $\frac{\partial \Pi_C}{\partial F_x(\widehat{x}_U)} = \rho(1-q)n\binom{n}{(1-q)n}(\rho F_x(\widehat{x}_U))^{(1-q)n-1}(1-\rho F_x(\widehat{x}_U))^{(1-q)n-1}\Big((1-\rho F_x(\widehat{x}_U))^{(2q-1)n+1} - (\rho F_x(\widehat{x}_U))^{(2q-1)n+1}\Big)$.

Combining everything together, we have $\frac{\partial \Pi_C(U)}{\partial \mu_x} = \phi(\frac{\widehat{x}_U - \mu_x}{\sigma_x})\frac{1}{\sigma_x}(\frac{\partial \widehat{x}_U}{\partial \mu_x} -$

$1)\rho(1-q)n\binom{n}{(1-q)n}(\rho F_x(\widehat{x}_U))^{(1-q)n-1}(1-\rho F_x(\widehat{x}_U))^{(1-q)n-1}\Big((1-\rho F_x(\widehat{x}_U))^{(2q-1)n+1}-(\rho F_x(\widehat{x}_U))^{(2q-1)n+1}\Big)$

Moreover, from Lemma 2.5.2, we also know that

$$\frac{\partial \widehat{x}_U}{\partial \mu_x} = \frac{\frac{(2q-1)n}{\sigma_x F_x(\widehat{x}_U)(1-\rho F_x(\widehat{x}_U))}\left(\frac{1-\rho F_x(\widehat{x}_U)}{\rho F_x(\widehat{x}_U)}\right)^{(2q-1)n}\phi\left(\frac{\widehat{x}_U-\mu_x}{\sigma_x}\right)}{\frac{(2q-1)n}{\sigma_x F_x(\widehat{x}_U)(1-\rho F_x(\widehat{x}_U))}\left(\frac{1-\rho F_x(\widehat{x}_U)}{\rho F_x(\widehat{x}_U)}\right)^{(2q-1)n}\phi\left(\frac{\widehat{x}_U-\mu_x}{\sigma_x}\right)+\frac{2(C-L)(R-L)}{(R-C)(R+C-2\widehat{x}_U)^2}} < 1$$

Hence we can conclude that $\frac{\partial \Pi_C(U)}{\partial \mu_x} > 0$ if and only if $t(U) > \frac{1}{2}$. ■

Proof of Theorem 2.5.6. In order to provide a proof for the statement of the lemma, we first consider an increase ϵ in q such that $(1-q)n$ changes by one. This implies that $\epsilon = \frac{1}{n}$. Then, we divide the change by ϵ. Hence, we define the change in $\Pi_C(U, \widehat{x}_U)$ as

$$\frac{d\Pi_C(U,\widehat{x}_U)}{dq} = n\Big(\frac{1}{n}\frac{\partial \Pi_C(U,\widehat{x}_U)}{\partial F_x(\widehat{x}_U)} f_x(\widehat{x}_U)\frac{\partial \widehat{x}_U}{\partial q} + \Pi_C(U,\widehat{x}_U, q+\frac{1}{n}) - \Pi_C(U,\widehat{x}_U,q)\Big)$$

In Lemma 2.5.3, we derived that $\frac{\partial \Pi_C}{\partial F_x(\widehat{x}_U)} = \rho(1-q)n\binom{n}{(1-q)n}(\rho F_x(\widehat{x}_U))^{(1-q)n-1}(1-\rho F_x(\widehat{x}_U))^{(1-q)n-1}\Big((1-\rho F_x(\widehat{x}_U))^{(2q-1)n+1}-(\rho F_x(\widehat{x}_U))^{(2q-1)n+1}\Big)$.

In order to figure out the direct effect of the change in q, keeping $\widehat{x}_U$ constant, an increase in q by $\frac{1}{n}$ only causes two additional terms to the summation, one preceding $t = (1-q)n$ and another one succeeding $t = qn$. Hence we have

$$\begin{aligned}\Pi_C(U,\widehat{x}_U,q+\frac{1}{n}) - \Pi_C(U,\widehat{x}_U,q) = &\binom{n}{qn+1}(\rho F_x(\widehat{x}_U))^{qn+1}(1-\rho F_x(\widehat{x}_U))^{(1-q)n-1}\\ &+\binom{n}{(1-q)n-1}(\rho F_x(\widehat{x}_U))^{(1-q)n-1}(1-\rho F_x(\widehat{x}_U))^{qn+1}\end{aligned}$$

Notice that

$$\binom{n}{qn+1} = \binom{n}{(1-q)n-1} = \frac{(1-q)n}{qn+1}\binom{n}{(1-q)n}$$

Hence we get

$$\begin{aligned}
&\Pi_C(U,\widehat{x}_U,q+\frac{1}{n})-\Pi_C(U,\widehat{x}_U,q)\\
&=\frac{(1-q)n}{qn+1}\binom{n}{(1-q)n}(\rho F_x(\widehat{x}_U)^{(1-q)n-1}(1-\rho F_x(\widehat{x}_U))^{(1-q)n-1}\\
&\Big((\rho F_x(\widehat{x}_U))^{(2q-1)n+2}+(1-\rho F_x(\widehat{x}_U))^{(2q-1)n+2}\Big)
\end{aligned}$$

Combining everything, we have

$$\begin{aligned}
\frac{d\Pi_C(U,\widehat{x}_U)}{dq}=&\frac{\partial\Pi_C(U,\widehat{x}_U)}{\partial F_x(\widehat{x}_U)}f_x(\widehat{x}_U)\frac{\partial\widehat{x}_U}{\partial q}+\frac{\Pi_C(U,\widehat{x}_U,q+\frac{1}{n})-\Pi_C(U,\widehat{x}_U,q)}{\frac{1}{n}}\\
=&\rho(1-q)n\binom{n}{(1-q)n}(\rho F_x(\widehat{x}_U))^{(1-q)n-1}(1-\rho F_x(\widehat{x}_U))^{(1-q)n-1}\\
&\Big((1-\rho F_x(\widehat{x}_U))^{(2q-1)n+1}-(\rho F_x(\widehat{x}_U))^{(2q-1)n+1}\Big)f_x(\widehat{x}_U)\frac{\partial\widehat{x}_U}{\partial q}\\
&+\frac{(1-q)n^2}{qn+1}\binom{n}{(1-q)n}(\rho F_x(\widehat{x}_U)^{(1-q)n-1}(1-\rho F_x(\widehat{x}_U))^{(1-q)n-1}\\
&\Big((\rho F_x(\widehat{x}_U))^{(2q-1)n+2}+(1-\rho F_x(\widehat{x}_U))^{(2q-1)n+2}\Big)\\
=&(1-q)n\binom{n}{(1-q)n}(\rho F_x(\widehat{x}_U))^{(1-q)n-1}(1-\rho F_x(\widehat{x}_U))^{(1-q)n-1}\\
&\Big[\rho f_x(\widehat{x}_U)\frac{\partial\widehat{x}_U}{\partial q}\Big((1-\rho F_x(\widehat{x}_U))^{(2q-1)n+1}-(\rho F_x(\widehat{x}_U))^{(2q-1)n+1}\Big)\\
&+\frac{1}{q+\frac{1}{n}}\Big((\rho F_x(\widehat{x}_U))^{(2q-1)n+2}+(1-\rho F_x(\widehat{x}_U))^{(2q-1)n+2}\Big)\Big]
\end{aligned}$$

Remember that $\frac{\partial\widehat{x}_U}{\partial q}>0$ if and only if $\rho F(\widehat{x}_U)<\frac{1}{2}$. Since $\Big((1-\rho F_x(\widehat{x}_U))^{(2q-1)n+1}-(\rho F_x(\widehat{x}_U))^{(2q-1)n+1}\Big)>0$ if and only if $\rho F(\widehat{x}_U)<\frac{1}{2}$ as well, we conclude that $\frac{d\Pi_C(U,\widehat{x}_U)}{dq}>0$. ■

Proof of Lemma 2.5.7. Since a change in C does not affect the limit values in the sum, the only effect of C is through the change in $\widehat{x}_U$. Hence we have

$$\frac{\partial\Pi_C(U,\widehat{x}_U)}{\partial C}=\frac{\partial\Pi_C(U,\widehat{x}_U)}{\partial F_x(\widehat{x}_U)}f_x(\widehat{x}_U)\frac{\partial\widehat{x}_U}{\partial C}$$

From Lemma 2.5.3, we have

$$\frac{\partial \Pi_C(U,\widehat{x}_U)}{\partial F_x(\widehat{x}_U)} = \rho(1-q)n\binom{n}{(1-q)n}(\rho F_x(\widehat{x}_U))^{(1-q)n-1}(1-\rho F_x(\widehat{x}_U))^{(1-q)n-1}$$
$$\Big((1-\rho F_x(\widehat{x}_U))^{(2q-1)n+1} - (\rho F_x(\widehat{x}_U))^{(2q-1)n+1}\Big)$$

which is positive if and only if $\rho F_x(\widehat{x}_U) < \frac{1}{2}$. Plugging in, we get

$$\frac{\partial \Pi_C(U,\widehat{x}_U)}{\partial C} = \rho(1-q)n\binom{n}{(1-q)n}(\rho F_x(\widehat{x}_U))^{(1-q)n-1}(1-\rho F_x(\widehat{x}_U))^{(1-q)n-1}$$
$$\Big((1-\rho F_x(\widehat{x}_U))^{(2q-1)n+1} - (\rho F_x(\widehat{x}_U))^{(2q-1)n+1}\Big) f_x(\widehat{x}_U)\frac{\partial \widehat{x}_U}{\partial C}$$

Hence, given the conditions for the sign of $\frac{\partial \widehat{x}_U}{\partial C}$ from Lemma 2.5.2, when $t(U) < \frac{1}{2}$, $\frac{\partial \Pi_C(U,\widehat{x}_U)}{\partial C} > 0$ if and only if $\widehat{x}_U > C$. When $t(U) > \frac{1}{2}$, $\frac{\partial \Pi_C(U,\widehat{x}_U)}{\partial C} > 0$ if and only if $\widehat{x}_U > C$ again, which concludes the proof. ■

Proof of Lemma 2.5.8. (i) Similar to the proof of Theorem 2.5.6 , we know that

$$\frac{d\Pi_R(U,\widehat{x}_U)}{dq} = n\Big(\frac{1}{n}\frac{\partial \Pi_R(U,\widehat{x}_U)}{\partial F_x(\widehat{x}_U)} f_x(\widehat{x}_U)\frac{\partial \widehat{x}_U}{\partial q} + \Pi_R(U,\widehat{x}_U,q+\frac{1}{n}) - \Pi_R(U,\widehat{x}_U,q)\Big)$$

and from Lemma 2.5.3 we have

$$\begin{aligned}&\frac{\partial}{\partial F_x(\widehat{x}_U)} P(T \leq (1-q)n-1 | F_x(\widehat{x}_U))\\ &= -\rho(1-q)n\binom{n}{(1-q)n}(\rho F_x(\widehat{x}_U))^{(1-q)n-1}(1-\rho F_x(\widehat{x}_U))^{qn}\end{aligned}$$

In order to figure out the direct effect of the change in q, keeping $\widehat{x}_U$ constant, an increase in q by $\frac{1}{n}$ only causes one term to drop from the summation, which is the one indexed $t = (1-q)n-1$. Hence we have

$$\begin{aligned}&\Pi_R(U,\widehat{x}_U,q+\frac{1}{n}) - \Pi_R(U,\widehat{x}_U,q)\\ &= -\binom{n}{(1-q)n-1}(\rho F_x(\widehat{x}_U))^{(1-q)n-1}(1-\rho F_x(\widehat{x}_U))^{qn+1}\end{aligned}$$

Notice that

$$\binom{n}{(1-q)n-1} = \frac{n!}{((1-q)n-1)!\,(qn+1)!} = \frac{(1-q)n}{qn+1}\binom{n}{(1-q)n}$$

Hence we get

$$\begin{aligned}
&\Pi_R(U, \widehat{x}_U, q+\frac{1}{n}) - \Pi_R(U, \widehat{x}_U, q)\\
&= -\frac{(1-q)n}{qn+1}\binom{n}{(1-q)n}(\rho F_x(\widehat{x}_U))^{(1-q)n-1}(1-\rho F_x(\widehat{x}_U))^{qn+1}
\end{aligned}$$

Combining everything, we have

$$\begin{aligned}
\frac{d\Pi_R(U,\widehat{x}_U)}{dq} &= \frac{\partial \Pi_R(U,\widehat{x}_U)}{\partial F_x(\widehat{x}_U)} f_x(\widehat{x}_U)\frac{\partial \widehat{x}_U}{\partial q} + \frac{\Pi_R(U,\widehat{x}_U,q+\frac{1}{n}) - \Pi_R(U,\widehat{x}_U,q)}{\frac{1}{n}}\\
=& -\rho(1-q)n\binom{n}{(1-q)n}(\rho F_x(\widehat{x}_U))^{(1-q)n-1}(1-\rho F_x(\widehat{x}_U))^{qn} f_x(\widehat{x}_U)\frac{\partial \widehat{x}_U}{\partial q}\\
-&\frac{(1-q)n^2}{qn+1}\binom{n}{(1-q)n}(\rho F_x(\widehat{x}_U))^{(1-q)n-1}(1-\rho F_x(\widehat{x}_U))^{qn+1}\\
=&(1-q)n\binom{n}{(1-q)n}(\rho F_x(\widehat{x}_U))^{(1-q)n-1}(1-\rho F_x(\widehat{x}_U))^{qn}\\
&\Big[-\rho f_x(\widehat{x}_U)\frac{\partial \widehat{x}_U}{\partial q} - \frac{1-\rho F_x(\widehat{x}_U)}{q+\frac{1}{n}}\Big]
\end{aligned}$$

This means that $\frac{d\Pi_R(U,\widehat{x}_U)}{dq} > 0$ if and only if

$$\frac{\partial \widehat{x}_U}{\partial q} < -\frac{1-\rho F_x(\widehat{x}_U)}{\rho f_x(\widehat{x}_U)(q+\frac{1}{n})}$$

where the right hand side of the inequality is negative. Moreover, we previously derived

$$\frac{\partial \widehat{x}_U}{\partial q} = \frac{(\frac{1-\rho F_x(\widehat{x}_U)}{\rho F_x(\widehat{x}_U)})^{(2q-1)n}}{\frac{2q-1}{2}\frac{f_x(\widehat{x}_U)}{F_x(\widehat{x}_U)(1-\rho F_x(\widehat{x}_U))}(\frac{1-\rho F_x(\widehat{x}_U)}{\rho F_x(\widehat{x}_U)})^{(2q-1)n} + \frac{(C-L)(R-L)}{(R-C)(R+C-2\widehat{x}_U)^2}\frac{1}{n}}\log(\frac{1-\rho F_x(\widehat{x}_U)}{\rho F_x(\widehat{x}_U)})$$

Hence we can conclude that $\frac{d\Pi_R(U,\widehat{x}_U)}{dq} < 0$ when $\rho F_x(\widehat{x}_U) < \frac{1}{2}$. *(ii)*Similarly, we know that

$$\frac{d\Pi_L(U,\widehat{x}_U)}{dq} = n\Big(\frac{1}{n}\frac{\partial \Pi_L(U,\widehat{x}_U)}{\partial F_x(\widehat{x}_U)} f_x(\widehat{x}_U)\frac{\partial \widehat{x}_U}{\partial q} + \Pi_L(U,\widehat{x}_U, q+\frac{1}{n}) - \Pi_L(U,\widehat{x}_U,q)\Big)$$

and from Lemma 2.5.3 we have

$$\frac{\partial}{\partial F_x(\widehat{x}_U)} P(T \geq qn+1 | F_x(\widehat{x}_U)) = \rho(qn+1)\binom{n}{qn+1}(\rho F_x(\widehat{x}_U))^{qn}(1-\rho F_x(\widehat{x}_U))^{(1-q)n-1}$$

In order to figure out the direct effect of the change in q, keeping $\widehat{x}_U$ constant, an increase in q by $\frac{1}{n}$ only causes one term to drop from the summation, which is the one indexed $t = qn+1$. Hence we have

$$\Pi_L(U,\widehat{x}_U, q+\frac{1}{n}) - \Pi_L(U,\widehat{x}_U,q) = -\binom{n}{qn+1}(\rho F_x(\widehat{x}_U))^{qn+1}(1-\rho F_x(\widehat{x}_U))^{(1-q)n-1}$$

Combining everything, we have

$$\begin{aligned}\frac{d\Pi_L(U,\widehat{x}_U)}{dq} =& \frac{\partial \Pi_L(U,\widehat{x}_U)}{\partial F_x(\widehat{x}_U)} f_x(\widehat{x}_U)\frac{\partial \widehat{x}_U}{\partial q} + \frac{\Pi_L(U,\widehat{x}_U, q+\frac{1}{n}) - \Pi_L(U,\widehat{x}_U,q)}{\frac{1}{n}} \\ =&\rho(qn+1)\binom{n}{qn+1}(\rho F_x(\widehat{x}_U))^{qn}(1-\rho F_x(\widehat{x}_U))^{(1-q)n-1} f_x(\widehat{x}_U)\frac{\partial \widehat{x}_U}{\partial q} \\ &-n\binom{n}{qn+1}(\rho F_x(\widehat{x}_U))^{qn+1}(1-\rho F_x(\widehat{x}_U))^{(1-q)n-1} \\ =&n\binom{n}{qn+1}(\rho F_x(\widehat{x}_U))^{qn}(1-\rho F_x(\widehat{x}_U))^{(1-q)n-1} \\ &\Big[(q+\frac{1}{n})\rho f_x(\widehat{x}_U)\frac{\partial \widehat{x}_U}{\partial q} - \rho F_x(\widehat{x}_U)\Big]\end{aligned}$$

This means that $\frac{d\Pi_L(U,\widehat{x}_U)}{dq} > 0$ if and only if

$$\frac{\partial \widehat{x}_U}{\partial q} > \frac{\rho F_x(\widehat{x}_U)}{(q+\frac{1}{n})\rho f_x(\widehat{x}_U)}$$

where the right hand side of the inequality is positive. Moreover, we previously derived

$$\frac{\partial \widehat{x}_U}{\partial q} = \frac{(\frac{1-\rho F_x(\widehat{x}_U)}{\rho F_x(\widehat{x}_U)})^{(2q-1)n}}{\frac{2q-1}{2}\frac{f_x(\widehat{x}_U)}{F_x(\widehat{x}_U)(1-\rho F_x(\widehat{x}_U))}(\frac{1-\rho F_x(\widehat{x}_U)}{\rho F_x(\widehat{x}_U)})^{(2q-1)n} + \frac{(C-L)(R-L)}{(R-C)(R+C-2\widehat{x}_U)^2}\frac{1}{n}} \log(\frac{1-\rho F_x(\widehat{x}_U)}{\rho F_x(\widehat{x}_U)})$$

Hence we can conclude that $\frac{d\Pi_L(U,\widehat{x}_U)}{dq} < 0$ when $\rho F_x(\widehat{x}_U) > \frac{1}{2}$. ■

Proof. In order to investigate the effects of a change in q on an individual voter from τ_L, we take the derivative of the expected utility from participating in the election with respect to q, which leads to

$$\begin{aligned}
&\frac{d}{dq}\mathbb{E}[u_L(j,U)] \\
&= -\Big[\frac{d\Pi_C(U,\widehat{x}_U)}{dq}(x_U - C)^2 + \frac{d\Pi_R(U,\widehat{x}_U)}{dq}(x_U - R)^2 + \frac{d\Pi_L(U,\widehat{x}_U)}{dq}(x_U - L)^2\Big]
\end{aligned}$$

We previously found that

$$\begin{aligned}
\frac{d\Pi_L(U,\widehat{x}_U)}{dq} =& \frac{(1-q)n}{qn+1}\binom{n}{(1-q)n}(\rho F_x(\widehat{x}_U))^{qn}(1-\rho F_x(\widehat{x}_U))^{(1-q)n-1} \\
&\Big[(qn+1)\rho f_x(\widehat{x}_U)\frac{\partial \widehat{x}_U}{\partial q} - n\rho F_x(\widehat{x}_U)\Big]
\end{aligned}$$

$$\begin{aligned}
\frac{d\Pi_R(U,\widehat{x}_U)}{dq} =& \frac{(1-q)n}{qn+1}\binom{n}{(1-q)n}(\rho F_x(\widehat{x}_U))^{(1-q)n-1}(1-\rho F_x(\widehat{x}_U))^{qn} \\
&\Big[-(qn+1)\rho f_x(\widehat{x}_U)\frac{\partial \widehat{x}_U}{\partial q} - n(1-\rho F_x(\widehat{x}_U))\Big]
\end{aligned}$$

and

$$\frac{d\Pi_C(U,\widehat{x}_U)}{dq} = \frac{(1-q)n}{qn+1}\binom{n}{(1-q)n}(\rho F_x(\widehat{x}_U))^{(1-q)n-1}(1-\rho F_x(\widehat{x}_U))^{(1-q)n-1}$$
$$\Big[\rho(qn+1)f_x(\widehat{x}_U)\frac{\partial \widehat{x}_U}{\partial q}\Big((1-\rho F_x(\widehat{x}_U))^{(2q-1)n+1}-(\rho F_x(\widehat{x}_U))^{(2q-1)n+1}\Big)$$
$$+\rho n\Big((\rho F_x(\widehat{x}_U))^{(2q-1)n+2}+(1-\rho F_x(\widehat{x}_U))^{(2q-1)n+2}\Big)\Big]$$

Plugging in, we get

$$\frac{d}{dq}\mathbb{E}[u_L(j,U)] = -\Bigg[\frac{(1-q)n}{qn+1}\binom{n}{(1-q)n}(\rho F_x(\widehat{x}_U))^{(1-q)n-1}(1-\rho F_x(\widehat{x}_U))^{(1-q)n-1}$$
$$\Big[\rho(qn+1)f_x(\widehat{x}_U)\frac{\partial \widehat{x}_U}{\partial q}\Big((1-\rho F_x(\widehat{x}_U))^{(2q-1)n+1}-(\rho F_x(\widehat{x}_U))^{(2q-1)n+1}\Big)$$
$$+\rho n\Big((\rho F_x(\widehat{x}_U))^{(2q-1)n+2}+(1-\rho F_x(\widehat{x}_U))^{(2q-1)n+2}\Big)\Big](x_{i,U}-C)^2$$
$$+\frac{(1-q)n}{qn+1}\binom{n}{(1-q)n}(\rho F_x(\widehat{x}_U))^{qn}(1-\rho F_x(\widehat{x}_U))^{(1-q)n-1}$$
$$\Big[(qn+1)\rho f_x(\widehat{x}_U)\frac{\partial \widehat{x}_U}{\partial q}-n\rho F_x(\widehat{x}_U)\Big](x_{i,U}-L)^2$$
$$+\frac{(1-q)n}{qn+1}\binom{n}{(1-q)n}(\rho F_x(\widehat{x}_U))^{(1-q)n-1}(1-\rho F_x(\widehat{x}_U))^{qn}$$
$$\Big[-(qn+1)\rho f_x(\widehat{x}_U)\frac{\partial \widehat{x}_U}{\partial q}-n(1-\rho F_x(\widehat{x}_U))\Big](x_{i,U}-R)^2\Bigg]$$

Rearranging we get

$$\frac{d}{dq}\mathbb{E}[u_L(j,U)] = -\frac{(1-q)n}{qn+1}\binom{n}{(1-q)n}(\rho F_x(\widehat{x}_U))^{(1-q)n-1}(1-\rho F_x(\widehat{x}_U))^{(1-q)n-1}$$

$$\Bigg[\Big[\rho(qn+1)f_x(\widehat{x}_U)\frac{\partial \widehat{x}_U}{\partial q}\Big((1-\rho F_x(\widehat{x}_U))^{(2q-1)n+1} - (\rho F_x(\widehat{x}_U))^{(2q-1)n+1}\Big)$$

$$+ \rho n\Big((\rho F_x(\widehat{x}_U))^{(2q-1)n+2} + (1-\rho F_x(\widehat{x}_U))^{(2q-1)n+2}\Big)\Big](x_{i,U} - C)^2$$

$$+ (\rho F_x(\widehat{x}_U))^{(2q-1)n+1}\Big[(qn+1)\rho f_x(\widehat{x}_U)\frac{\partial \widehat{x}_U}{\partial q} - n\rho F_x(\widehat{x}_U)\Big](x_{i,U} - L)^2$$

$$+ (1-\rho F_x(\widehat{x}_U))^{(2q-1)n+1}$$

$$\Big[-(qn+1)\rho f_x(\widehat{x}_U)\frac{\partial \widehat{x}_U}{\partial q} - n(1-\rho F_x(\widehat{x}_U))\Big](x_{i,U} - R)^2\Bigg]$$

Further algebra shows that

$$\frac{d}{dq}\mathbb{E}[u_L(j,U)] = -\frac{\rho(1-q)n^2}{qn+1}\binom{n}{(1-q)n}(\rho F_x(\widehat{x}_U))^{(1-q)n-1}(1-\rho F_x(\widehat{x}_U))^{(1-q)n-1}$$

$$\Bigg[(q+\frac{1}{n})f_x(\widehat{x}_U)\frac{\partial \widehat{x}_U}{\partial q}(1-\rho F_x(\widehat{x}_U))^{(2q-1)n+1}[(x_{i,U} - C)^2 - (x_{i,U} - R)^2]$$

$$+ (q+\frac{1}{n})f_x(\widehat{x}_U)\frac{\partial \widehat{x}_U}{\partial q}(\rho F_x(\widehat{x}_U))^{(2q-1)n+1}[(x_{i,U} - L)^2 - (x_{i,U} - C)^2]$$

$$+ (\rho F_x(\widehat{x}_U))^{(2q-1)n+2}[(x_{i,U} - C)^2 - (x_{i,U} - L)^2]$$

$$+ (1-\rho F_x(\widehat{x}_U))^{(2q-1)n+2}[(x_{i,U} - C)^2 - (x_{i,U} - R)^2]\Bigg]$$

Rearranging again we get

$$\frac{d}{dq}\mathbb{E}[u_L(j,U)] = -\frac{\rho(1-q)n^2}{qn+1}\binom{n}{(1-q)n}(\rho F_x(\widehat{x}_U))^{(1-q)n-1}(1-\rho F_x(\widehat{x}_U))^{(1-q)n-1}$$

$$\Bigg[(1-\rho F_x(\widehat{x}_U))^{(2q-1)n+1}[(q+\frac{1}{n})f_x(\widehat{x}_U)\frac{\partial \widehat{x}_U}{\partial q} + 1 - \rho F_x(\widehat{x}_U)]$$

$$[(x_{i,U} - C)^2 - (x_{i,U} - R)^2] + (\rho F_x(\widehat{x}_U))^{(2q-1)n+1}[(q+\frac{1}{n})f_x(\widehat{x}_U)\frac{\partial \widehat{x}_U}{\partial q}$$

$$- \rho F_x(\widehat{x}_U)][(x_{i,U} - L)^2 - (x_{i,U} - C)^2]\Bigg]$$

Hence, for a voter with preference parameter $x_{i,U}$, $\frac{d}{dq}\mathbb{E}[u_L(j,U)] > 0$ if and only if

$$(1-\rho F_x(\widehat{x}_U))^{(2q-1)n+1}[(q+\frac{1}{n})f_x(\widehat{x}_U)\frac{\partial \widehat{x}_U}{\partial q}+1-\rho F_x(\widehat{x}_U)]$$
$$[(x_{i,U}-C)^2-(x_{i,U}-R)^2]+$$
$$(\rho F_x(\widehat{x}_U))^{(2q-1)n+1}[(q+\frac{1}{n})f_x(\widehat{x}_U)\frac{\partial \widehat{x}_U}{\partial q}-\rho F_x(\widehat{x}_U)]$$
$$[(x_{i,U}-L)^2-(x_{i,U}-C)^2]<0$$

which holds if and only if

$$\frac{(x_{i,U}-L)^2-(x_{i,U}-C)^2}{(x_{i,U}-C)^2-(x_{i,U}-R)^2} < -\frac{(1-\rho F_x(\widehat{x}_U))^{(2q-1)n+1}[(q+\frac{1}{n})f_x(\widehat{x}_U)\frac{\partial \widehat{x}_U}{\partial q}+1-\rho F_x(\widehat{x}_U)]}{(\rho F_x(\widehat{x}_U))^{(2q-1)n+1}[(q+\frac{1}{n})f_x(\widehat{x}_U)\frac{\partial \widehat{x}_U}{\partial q}-\rho F_x(\widehat{x}_U)]}$$

for $x_{i,U} < \frac{R+C}{2}$ and $(q+\frac{1}{n})f_x(\widehat{x}_U)\frac{\partial \widehat{x}_U}{\partial q}-\rho F_x(\widehat{x}_U) > 0$. Moreover, notice that $\frac{(x_{i,U}-L)^2-(x_{i,U}-C)^2}{(x_{i,U}-C)^2-(x_{i,U}-R)^2} = -\frac{(2x_{i,U}-C-L)(C-L)}{(R+C-2x_{i,U})(R-C)}$, which is the left hand side of inequality 2.2. Hence we can rewrite the inequality above as

$$\frac{(2x_{i,U}-C-L)(C-L)}{(R+C-2x_{i,U})(R-C)} > \frac{(1-\rho F_x(\widehat{x}_U))^{(2q-1)n+1}[(q+\frac{1}{n})f_x(\widehat{x}_U)\frac{\partial \widehat{x}_U}{\partial q}+1-\rho F_x(\widehat{x}_U)]}{(\rho F_x(\widehat{x}_U))^{(2q-1)n+1}[(q+\frac{1}{n})f_x(\widehat{x}_U)\frac{\partial \widehat{x}_U}{\partial q}-\rho F_x(\widehat{x}_U)]}$$

again for $x_{i,U} < \frac{R+C}{2}$ and $(q+\frac{1}{n})f_x(\widehat{x}_U)\frac{\partial \widehat{x}_U}{\partial q}-\rho F_x(\widehat{x}_U) > 0$. Combining this with inequality 2.4, we get the condition for being better off for a τ_L agent who is voting l (who has $x_{i,U} < \frac{R+C}{2}$ by definition)

$$\frac{\rho F_x(\widehat{x}_U)}{1-\rho F_x(\widehat{x}_U)} > \frac{(q+\frac{1}{n})f_x(\widehat{x}_U)\frac{\partial \widehat{x}_U}{\partial q}+1-\rho F_x(\widehat{x}_U)}{(q+\frac{1}{n})f_x(\widehat{x}_U)\frac{\partial \widehat{x}_U}{\partial q}-\rho F_x(\widehat{x}_U)}$$

assuming $(q+\frac{1}{n})f_x(\widehat{x}_U)\frac{\partial \widehat{x}_U}{\partial q}-\rho F_x(\widehat{x}_U) > 0$.

From Lemma 2.5.2, we have

$$\frac{\partial \widehat{x}_U}{\partial q} = \frac{(\frac{1-\rho F_x(\widehat{x}_U)}{\rho F_x(\widehat{x}_U)})^{(2q-1)n}}{\frac{2q-1}{2}\frac{f_x(\widehat{x}_U)}{F_x(\widehat{x}_U)(1-\rho F_x(\widehat{x}_U))}(\frac{1-\rho F_x(\widehat{x}_U)}{\rho F_x(\widehat{x}_U)})^{(2q-1)n} + \frac{(C-L)(R-L)}{(R-C)(R+C-2\widehat{x}_U)^2}\frac{1}{n}} \log(\frac{1-\rho F_x(\widehat{x}_U)}{\rho F_x(\widehat{x}_U)})$$

which means that $(q+\frac{1}{n})f_x(\widehat{x}_U)\frac{\partial \widehat{x}_U}{\partial q} - \rho F_x(\widehat{x}_U) > 0$ holds if and only if

$$\frac{(q+\frac{1}{n})(\frac{1-\rho F_x(\widehat{x}_U)}{\rho F_x(\widehat{x}_U)})^{(2q-1)n}}{\frac{q-\frac{1}{2}}{\rho F_x(\widehat{x}_U)^2(1-\rho F_x(\widehat{x}_U))}(\frac{1-\rho F_x(\widehat{x}_U)}{\rho F_x(\widehat{x}_U)})^{(2q-1)n} + \frac{(C-L)(R-L)}{(R-C)(R+C-2\widehat{x}_U)^2}\frac{1}{n}} \log(\frac{1-\rho F_x(\widehat{x}_U)}{\rho F_x(\widehat{x}_U)}) > 1$$

Notice that this could happen only if $t(U) < \frac{1}{2}$. But then the left hand side of the condition for being better off for a τ_L agent who is voting l is less than 1 and right hand side is greater than 1, which is a contradiction. So we can conclude that no τ_L agent who is voting l can be better off as a result of an increase in q when $(q+\frac{1}{n})f_x(\widehat{x}_U)\frac{\partial \widehat{x}_U}{\partial q} - \rho F_x(\widehat{x}_U) > 0$.

Now assume $(q+\frac{1}{n})f_x(\widehat{x}_U)\frac{\partial \widehat{x}_U}{\partial q} - \rho F_x(\widehat{x}_U) < 0$. This means that, for a voter with preference parameter $x_{i,U}$, $\frac{d}{dq}\mathbb{E}[u_L(j,U)] > 0$ if and only if

$$\frac{(2x_{i,U}-C-L)(C-L)}{(R+C-2x_{i,U})(R-C)} < \frac{(1-\rho F_x(\widehat{x}_U))^{(2q-1)n+1}[(q+\frac{1}{n})f_x(\widehat{x}_U)\frac{\partial \widehat{x}_U}{\partial q} + 1 - \rho F_x(\widehat{x}_U)]}{(\rho F_x(\widehat{x}_U))^{(2q-1)n+1}[(q+\frac{1}{n})f_x(\widehat{x}_U)\frac{\partial \widehat{x}_U}{\partial q} - \rho F_x(\widehat{x}_U)]}$$

Notice that again this is not possible since when $t(U) < \frac{1}{2}$, we have $\frac{\partial \widehat{x}_U}{\partial q} > 0$, which implies that $(q+\frac{1}{n})f_x(\widehat{x}_U)\frac{\partial \widehat{x}_U}{\partial q} - \rho F_x(\widehat{x}_U) > -1$ since $\rho F_x(\widehat{x}_U) < 1$. This means that the right hand side above is negative but the left hand side is positive, a contradiction to the inequality. Hence, we can conclude that if $t(U) < \frac{1}{2}$, the τ_L agent who are voting l gets worse off as a result of an increase in q.

Now assume $(q+\frac{1}{n})f_x(\widehat{x}_U)\frac{\partial \widehat{x}_U}{\partial q} - \rho F_x(\widehat{x}_U) < 0$ and $t(U) > \frac{1}{2}$ (since $(q+\frac{1}{n})f_x(\widehat{x}_U)\frac{\partial \widehat{x}_U}{\partial q} - \rho F_x(\widehat{x}_U) > 0$ and $t(U) > \frac{1}{2}$ cannot happen at the same time, this is the only case to consider). Again, this means that, for a voter with preference parameter $x_{i,U}$, $\frac{d}{dq}\mathbb{E}[u_L(j,U)] > 0$ if and only if

$$\frac{(2x_{i,U}-C-L)(C-L)}{(R+C-2x_{i,U})(R-C)} < \frac{(1-\rho F_x(\widehat{x}_U))^{(2q-1)n+1}[(q+\frac{1}{n})f_x(\widehat{x}_U)\frac{\partial \widehat{x}_U}{\partial q}+1-\rho F_x(\widehat{x}_U)]}{(\rho F_x(\widehat{x}_U))^{(2q-1)n+1}[(q+\frac{1}{n})f_x(\widehat{x}_U)\frac{\partial \widehat{x}_U}{\partial q}-\rho F_x(\widehat{x}_U)]}$$

and if $(q+\frac{1}{n})f_x(\widehat{x}_U)\frac{\partial \widehat{x}_U}{\partial q}-\rho F_x(\widehat{x}_U) > -1$, we reach to a contradiction. Now assume $(q+\frac{1}{n})f_x(\widehat{x}_U)\frac{\partial \widehat{x}_U}{\partial q}-\rho F_x(\widehat{x}_U) < -1$. The right hand side is positive and less than 1. Notice that for all τ_L voters voting l we have $0 < \frac{(2x_{i,U}-C-L)(C-L)}{(R+C-2x_{i,U})(R-C)} < \frac{(1-\rho F_x(\widehat{x}_U))^{(2q-1)n}}{(\rho F_x(\widehat{x}_U))^{(2q-1)n}} < 1$ when $t(U) > \frac{1}{2}$. Hence we cannot derive any conclusions without the specific values of the parameters. ■

Proof of Lemma 2.5.10. Convergence of $\widehat{x}_{U,n}$: For a given n, we have the following equation that pins down the cutoff value:

$$\Big(\frac{1-t(U,n)}{t(U,n)}\Big)^{(2q-1)n} = \frac{(2\widehat{x}_{U,n}-L-C)(C-L)}{(C+R-2\widehat{x}_{U,n})(R-C)}$$

Notice that $\lim_{n\to\infty}\left(\frac{1-t(U,n)}{t(U,n)}\right)^{(2q-1)n} \to \infty$ if $\lim_{n\to\infty} t(U,n) < \frac{1}{2}$, and $\lim_{n\to\infty}\left(\frac{1-t(U,n)}{t(U,n)}\right)^{(2q-1)n} = 0$ if $\lim_{n\to\infty} t(U,n) > \frac{1}{2}$. If $\lim_{n\to\infty} t(U,n) = \frac{1}{2}$, then $\lim_{n\to\infty}\left(\frac{1-t(U,n)}{t(U,n)}\right)^{(2q-1)n}$ might diverge to infinity or converge to a positive real number, depending on the functional form of $t(U,n)$.

First assume that $F_x(\frac{C+L}{2}) > \frac{1}{2\rho}$. This means that $t(U,n) > \frac{1}{2}\ \forall n$. Then $\lim_{n\to\infty}\left(\frac{1-t(U,n)}{t(U,n)}\right)^{(2q-1)n} = 0$, which implies that the cutoff is $\frac{C+L}{2}$, according to the equation above. Now assume that $F_x(\frac{R+C}{2}) < \frac{1}{2\rho}$. This means that $t(U,n) < \frac{1}{2}\ \forall n$. Then $\lim_{n\to\infty}\left(\frac{1-t(U,n)}{t(U,n)}\right)^{(2q-1)n} \to \infty$, which implies that the cutoff converges to $\frac{R+C}{2}$, according to the equation above. Finally, assume that $F_x(\frac{C+L}{2}) < \frac{1}{2\rho} < F_x(\frac{R+C}{2})$. The lemma above implies that $\widehat{x}_{U,n}$ is a monotonic sequence on a bounded space, which means that it converges to some $\overline{x}_U \in (\frac{C+L}{2}, \frac{R+C}{2})$. For this $\overline{x}_U$ to constitute an equilibrium of the game in the limit, the following has to hold:

$$\lim_{n\to\infty}\Big(\frac{1-t(U,n)}{t(U,n)}\Big)^{(2q-1)n}=\frac{(2\overline{x}_U-L-C)(C-L)}{(C+R-2\overline{x}_U)(R-C)}$$

which holds if and only if $\lim_{n\to\infty} t(U,n)=\frac{1}{2}$. This means that $\lim_{n\to\infty} t(U)=\frac{1}{2}$ when $x_U=\overline{x}_U$. Hence we have $\overline{x}_U=F^{-1}(\frac{1}{2\rho})$, which is in $(\frac{C+L}{2},\frac{R+C}{2})$. Moreover, notice that

$$t(U,n)=\frac{1}{1+\Big(\frac{(2F^{-1}(\frac{1}{2\rho})-L-C)(C-L)}{(C+R-2F^{-1}(\frac{1}{2\rho}))(R-C)}\Big)^{\frac{1}{(2q-1)n}}}$$

gives a functional form for one such $t(U,n)$. Even though the functional form of $t(U,n)$ is not unique, they all converge to $\frac{1}{2}$.

Convergence of $\widehat{y}_{D,n}$: An argument similar to the one above gives the desired results.

The derivation of $t(\omega,\overline{\pi})$ is straightforward. ■

Proof of Lemma 2.5.11. Assume the state is U. By assumption, $q>\frac{1}{2}>1-q$. L is implemented if and only if either $\rho F_x(\frac{C+L}{2})>q$ and $F_x(\frac{C+L}{2})>\frac{1}{2\rho}$ or $\rho F_x(\frac{R+C}{2})>q$ and $F_x(\frac{R+C}{2})<\frac{1}{2\rho}$. Notice neither can hold because $F_x(x)<1$ $\forall x$. Hence L cannot be implemented. R can be implemented if and only if either $\rho F_x(\frac{C+L}{2})<1-q$ and $F_x(\frac{C+L}{2})>\frac{1}{2\rho}$ or $\rho F_x(\frac{R+C}{2})<1-q$ and $F_x(\frac{R+C}{2})<\frac{1}{2\rho}$. First assume $\rho F_x(\frac{C+L}{2})<1-q$ and $F_x(\frac{C+L}{2})>\frac{1}{2\rho}$. These imply $\frac{1-q}{\rho}>F_x(\frac{C+L}{2})>\frac{1}{2\rho}$, but then $1-q<\frac{1}{2}$, which is a contradiction. Now assume $\rho F_x(\frac{R+C}{2})<1-q$ and $F_x(\frac{R+C}{2})<\frac{1}{2\rho}$. The first inequality implies the second one, yielding the desired result.

A similar set of arguments gives the results for the state D. ■

Proof of 2.6.1. Existence of $\widehat{x}_{U,d}$ and $\widehat{x}_{U,u}$ comes from the fact that the right hand side of inequality 2.4 is continuous, positive for $\widehat{x}_{U,s}<\frac{C+L}{2}$ and diverges to negative infinity as $\widehat{x}_{U,s}\to\infty$, and the left hand side is -1. $\widehat{x}_{U,d}>\widehat{x}_{U,u}$ follows from the fact that $\frac{\beta(U|u)}{\beta(D|u)}>1>\frac{\beta(U|d)}{\beta(D|d)}$. A similar argument establishes the existence of $\widehat{y}_{D,u}<\widehat{y}_{D,d}$.

The construction of $t(\omega,\overline{\pi})$ is straightforward. ■

Proof of Theorem 2.6.3. Assume $F_x(\frac{R+C}{2}) < \frac{1-q}{\rho}$ and $F_y(\frac{C+L}{2}) \leq \frac{q-\rho}{1-\rho}$. This implies that when the state is known, R is implemented when the state is U and C is implemented in state D. Assume there is information aggregation, i.e., there exists a strategy profile that implements R when the state is U and C in state D in the elections with unknown state as well. In the rest of the proof, we will show that there always exist feasible F_x and F_y such that they satisfy the conditions above, but lead to the outcome C in state U for any sequence of equilibrium strategy profiles.

If R is implemented, that means that there exist two sequences of cutoffs $\widehat{x}_{U,u,n}$ and $\widehat{x}_{U,d,n}$ according to the lemma above, pinned down by

$$\frac{P(piv_R|D)}{P(piv_L|D)}\frac{R-C}{C-L}\frac{R+C-2x_D}{C+L-2x_D}+\frac{\beta(U|s)}{\beta(D|s)}\frac{P(piv_L|U)}{P(piv_L|D)}\frac{C+L-2\widehat{x}_{U,s,n}}{C+L-2x_D} \\ +\frac{\beta(U|s)}{\beta(D|s)}\frac{P(piv_R|U)}{P(piv_L|D)}\frac{R-C}{C-L}\frac{R+C-2\widehat{x}_{U,s,n}}{C+L-2x_D}=-1$$

The critical pivotality ratios are as follows: $\frac{P(piv_R|D)}{P(piv_L|D)} = \left(\frac{1-t(D)}{t(D)}\right)^{(2q-1)n}$, $\frac{P(piv_L|U)}{P(piv_L|D)} = \left(\frac{t(U)}{t(D)}\right)^{qn}\left(\frac{1-t(U)}{1-t(D)}\right)^{(1-q)n}$ and $\frac{P(piv_R|U)}{P(piv_L|D)} = \left(\frac{1-t(U)}{t(D)}\right)^{qn}\left(\frac{t(U)}{1-t(D)}\right)^{(1-q)n}$. Notice that if there is information aggregation, then it must be the case that $\lim_{n\to\infty} t(U) < 1-q$ and $q \geq \lim_{n\to\infty} t(D) \geq 1-q$ as $n\to\infty$ in equilibrium. This means that $\lim_{n\to\infty}\frac{P(piv_L|U)}{P(piv_L|D)} = 0$.

First, assume that $\lim_{n\to\infty} t(D) > \frac{1}{2}$. This implies $\lim_{n\to\infty}\frac{P(piv_R|D)}{P(piv_L|D)} = 0$. Then, the limit cutoff, or lack thereof, is determined by the third term on the equation above, where the critical pivotality ratio can take any value. First, suppose that $\lim_{n\to\infty}\frac{P(piv_R|U)}{P(piv_L|D)} = 0$. Then this means all τ_L vote l regardless of their preference parameter or their signal, a contradiction to $\lim_{n\to\infty} t(U) < 1-q$. Next, suppose $\lim_{n\to\infty}\frac{P(piv_R|U)}{P(piv_L|D)}$ converges to a strictly positive real number. This would imply that both of the cutoffs are higher than $\frac{R+C}{2}$. Then let $\overline{x}_{U,u,n} = \lim_{n\to\infty}\widehat{x}_{U,u,n}$. Construct F_x in such a way that $F_x(\frac{R+C}{2}) < \frac{1-q}{\rho}$ holds, but $F_x(\overline{x}_{U,u,n}) > \frac{1-q}{\rho}$, which is possible since $\overline{x}_{U,u,n} > \frac{R+C}{2}$. Lastly, suppose $\lim_{n\to\infty}\frac{P(piv_R|U)}{P(piv_L|D)} \to \infty$.

Then both cutoffs converge to $\frac{R+C}{2}$, which implies that all τ_L with $x_{U,i} < \frac{R+C}{2}$ vote l. Consider the indifference condition of the τ_R agents,

$$\frac{P(piv_L|U)}{P(piv_R|U)}\frac{C-L}{R-C}\frac{2y_U-C-L}{2y_U-R-C}+\frac{\beta(D|s)}{\beta(U|s)}\frac{P(piv_R|D)}{P(piv_R|U)}\frac{2y_{D,i}-R-C}{2y_U-R-C}$$
$$+\frac{\beta(D|s)}{\beta(U|s)}\frac{P(piv_L|D)}{P(piv_R|U)}\frac{C-L}{R-C}\frac{2y_{D,i}-C-L}{2y_U-R-C}\geq -1$$

where the critical pivotality ratios are $\frac{P(piv_L|U)}{P(piv_R|U)} = \left(\frac{t(U)}{1-t(U)}\right)^{(2q-1)n}$, $\frac{P(piv_R|D)}{P(piv_R|U)} = \left(\frac{t(D)}{t(U)}\right)^{(1-q)n}\left(\frac{1-t(D)}{1-t(U)}\right)^{qn}$ and $\frac{P(piv_L|D)}{P(piv_R|U)} = \left(\frac{t(D)}{1-t(U)}\right)^{qn}\left(\frac{1-t(D)}{t(U)}\right)^{(1-q)n}$. Notice that $\lim_{n\to\infty}\frac{P(piv_L|U)}{P(piv_R|U)} = 0$ by assumption and $\lim_{n\to\infty}\frac{P(piv_L|D)}{P(piv_R|U)} = 0$ by supposition. If $\lim_{n\to\infty}\frac{P(piv_R|D)}{P(piv_R|U)} = 0$, then all τ_R votes r regardless of their preference parameter or signal. Then, we would have $t(D) = \rho(1-F_x(\frac{R+C}{2}))+(1-\rho)$. But that contradicts $t(D) < q$. If $\lim_{n\to\infty}\frac{P(piv_R|D)}{P(piv_R|U)}$ is a positive real number, then both of the cutoffs are lower than $\frac{R+C}{2}$. Let $\overline{y}_{D,u} = \lim_{n\to\infty}\widehat{y}_{D,u,n}$. This implies all τ_R with $y_{D,i} < \overline{y}_{D,u}$ always vote l regardless of their signal. Then construct F_y in such a way that $\rho F_x(\frac{R+C}{2}) + (1-\rho)F_y(\overline{y}_{D,u}) > 1-q$, which is possible, yielding a contradiction to $t(U) < 1-q$. Lastly, if $\lim_{n\to\infty}\frac{P(piv_R|D)}{P(piv_R|U)} =\to \infty$, then all τ_R with $y_{D,i} < \frac{R+C}{2}$ votes l. Again, construct F_y in such a way that $\rho F_x(\frac{R+C}{2}) + (1-\rho)F_y(\frac{R+C}{2}) > 1-q$, which is possible, yielding a contradiction to $t(U) < 1-q$. This concludes that there is no sequence of equilibrium strategy profiles that leads to information aggregation and $\lim_{n\to\infty} t(D) > \frac{1}{2}$.

Now, assume that $\lim_{n\to\infty} t(D) < \frac{1}{2}$. This implies $\lim_{n\to\infty}\frac{P(piv_R|D)}{P(piv_L|D)} \to \infty$. Since the right hand side of the first equation is negative, if the left hand side diverges to positive infinity as well, we are done. If not, then all the arguments presented above still go through, since the first term is always positive. ■

Chapter 3

Protests, Strategic Information Provision and Political Communication

3.1 Introduction

Protests are a common component of political systems in many countries around the world. In many instances, protests serve as a tool for the citizens to convey their information to the incumbent, such as their sentiments about the regime and its performance. Hence, the incumbent can learn vital information from the protests that might be relevant to her decisions. However, the media can affect the citizens' decision to participate in these protests and how the policies are determined as a result of these protests (Walgrave & Vliegenthart, 2012), and sometimes the media might have its own goals (Amenta et al., 2017). For example, in the Vietnam War era in the United States, widespread protests around the country informed the policymakers regarding the negative sentiments of a sizable portion of the public about how the country handled its involvement in the conflict (Anderson, 2002). These protests had an undeniable role in the policymakers' decision to withdraw from the war later in 1973 (McAdam & Su, 2002). Experts believe that the media coverage of the war affected the protest participation and how the protests led to changes in policymaking (Gitlin, 2003). In general, there is extensive research that shows that media framing affect people's attitudes towards issues (Nelson et al., 1997) and more specifically their decision to participate in or support protests (Cooper, 2002; Brown & Mourão, 2021). Protests then have a direct or indirect effect on policy change (Madestam et al., 2013).

In this paper, we look at the familiar problem of protests as a political communication environment, but we add a biased information provider as an intermediary. We investigate the kind of information transmission schemes the biased information provider commits to, and we try to understand whether an incumbent can still learn important information from the protesters even though their information source is biased.

In order to investigate the relationship between protests, political communication and biased media, we set up a global games model as introduced by Carlsson & Van Damme (1993), which is commonly used in various contexts to theoretically understand political protest and so-called 'informal elections' (some examples are Edmond (2013), Persson & Tabellini (2009), Shadmehr & Bernhardt (2011), Boix & Svolik (2013), De Mesquita (2010)). In our model, there is a possible protest which, if successful, removes the incumbent from office. We assume that the protest success threshold is set by the incumbent prior to the protest. We incorporate biased media as an information provider who sends signals to the citizens regarding the payoff relevant state of the world, in order to facilitate or prevent a successful protest, depending on its type. We assume that the information provider can control how noisy the signals are.

In the first part of the paper, we focus on the interaction between the citizens and the strategic information provider, and assume that the protest success threshold is a predetermined parameter, i.e. we first consider a simpler model where the incumbent is not a strategic player. More specifically, we look at the problem of a biased information provider who is deciding on the precision of the signals it sends about the state of the world. These signals in turn affect the citizens' participation decisions. The strategic information provider is trying to either instigate a successful protest or dissuade the protest depending on its type, which is common knowledge. If the information provider is trying to dissuade the protest, we find that the information provider sends completely uninformative signals if the success threshold is more than half of the population, but there is an optimal level of noise if the threshold is less than half. In the case where the success threshold is more than half, the provider is able to successfully dissuade a protest by providing no information and forcing the citizens to rely on their priors since the threshold is already high. But if the threshold is relatively low, then there is a positive chance that the protest is successful if the citizens rely on their priors, hence there is an optimal level of noise which minimizes the probability that the protest is successful. If the success threshold is even lower, then the information provider perfectly reveals the state of the world in the hopes that it is low enough to prevent a successful protest. If the information provider is trying to instigate a protest, then we find that the information provider sends completely uninformative signals when the threshold is less than half, due to similar reasons explained above. But when threshold is more than half, then signals that perfectly reveal the state are optimal. Since a successful protest is unlikely when the success threshold is high, the best the provider can do is to commit to perfectly revealing the state of the world and hope that is high enough to lead to a successful protest. We also look at the effects of the equilibrium noise parameter decision of the information provider on the welfare of the citizens.

In the second part of the paper, we turn the incumbent into a strategic player as well. We assume that the incumbent's payoff from remaining in office depends on the state of the world but she is biased; she finds it beneficial to leave office if the state of the world is relatively high (compared to what would be welfare maximizing for the citizens). With this formulation, we investigate her choice of a success threshold for the protest. Our main aim with this second part of the paper is to understand if political communication is possible under such a setting, where it is meddled by a strategic information provider. In order to answer this question, we first look at the choice of success threshold by the incumbent, under different types of information providers. In the case where the information provider is trying to dissuade the protest, we find that if the bias is relatively low, then there exists an optimal success threshold that is less than half. Since the provider is trying to dissuade the protest, the incumbent needs to set a relatively lower threshold for a successful protest, since she is minimally biased. If the incumbent bias is at a moderate level, then the optimal threshold is half the population and if the bias is relatively high, then any threshold that is more than half is optimal, since all of them lead to a failed protest. If the information provider is trying to instigate a protest, then it turns out that there is a cutoff level of incumbent bias, under which the incumbent sets an optimal threshold as a function of her bias, and over which she sets the success threshold to 1 (i.e. she requires every citizen to participate in the protest to call it a successful protest), to minimize her chances of leaving office. Importantly, we find that, in equilibrium, the information provider who is trying to instigate a protest always perfectly reveals the state of the world, whereas the information provider who is trying to dissuade the protest sends either noisy or completely uninformative signals. We also look at the effects of the incumbent bias on the welfare of the citizens.

Lastly, we also comment on the political communication aspect. We say that political communication is possible if and only if the incumbent can learn about the state of the world by looking at the protest turnout. It turns out that if the information provider is trying to instigate a protest, then political communication is always possible, regardless of the level of bias of the incumbent. But, if the information provider is trying to dissuade the protest, then political communication is possible if and only if the incumbent bias is relatively small. We also briefly comment on the commitment assumption for the incumbent. More specifically, we show that, since the incumbent must commit to a protest success threshold ex-ante, she almost always disregards some of the informational content of the protest in equilibrium, in the sense that her actions do not match what her actions would have been had she known the state of the world beforehand. Hence, even under the incumbent bias levels that lead to political communication, she fails to replicate her perfect information actions, as a result of her ex-ante commitment. This result also speaks to the importance of the commitment assumption on driving our results.

The rest of the paper is structured as follows. Section 2 goes over the relevant literature on global games and protest from economics and theoretical political science. Section 3 introduces the base model without a strategic incumbent, and pins down the equilibrium behavior of the citizens and both types of information providers, along with presenting some comparative statics and welfare results. Section 4 builds on the model presented in Section 3 by incorporating a strategic incumbent and explores her equilibrium behavior and looks at the possibility of political communication. Section 5 contains concluding remarks. All proofs are relegated to the Appendix.

3.2 Related Literature

This work is strongly related to the global games literature, and the broad line of protests research within the domain of economics and theoretical political science. Morris & Shin (2001) provides an excellent survey into the foundations of global games.

Global games are first explored by Carlsson & Van Damme (1993) to study the incomplete information games where the payoffs are not deterministic. Morris & Shin (1998) employed the global games framework to study currency attacks, and later considered how public information can affect the coordination problems inherent in the framework (Morris & Shin, 2002).

Global games formulation of protest rose to prominence since it highlights the view of protests as a coordination problem. Shadmehr & Bernhardt (2011) provide a formal global games model of protests where the citizens are uncertain about the payoff of status quo and regime change, and investigates the coordination properties that it leads to. Persson & Tabellini (2009) use the global games framework to explore the dynamics of regime change and relationship between economic growth and democratic capital. Boix & Svolik (2013) consider a model that articulates how the threat of protests and regime change influence dictatorial power sharing. Edmond (2013) considers a global games formulation of protests where the incumbent leaves office if the protest is successful, and the incumbent can manipulate citizen belief through propaganda. They find that if the signal precision is high, then the incumbent if less likely to leave office in equilibrium. In a similar paper, De Mesquita (2010) considers how the use of violence can be used as a public belief manipulation tool. Even though the initial premise of this paper is similar to ours, our model differs because we consider an information provider who is manipulating beliefs, and the incumbent does not necessarily want to remain in office regardless of the state of the world. Little et al. (2012) is the paper that is the most similar in the literature to our work presented in this paper in term of the basis of the model. They present a global games model where the elections determine whether the incumbent stays in office or not but also affect public opinion, but the incumbent can also fraudulently affect the election results. Our formulation of protests essentially removes the elections from this paper, and turns the information provider and the incumbent into strategic agents. Hence, the orientations of the papers in terms of the results presented are completely different. Little et al. (2015) is also similar. They present a global games model where the elections determine whether the incumbent stays in office or not, but also the citizens can participate in the protests to remove the incumbent when the elections results are close.

There are other formulations of protests that do not use a global games framework. These papers employ an election framework, since protests can be considered an informal type of elections. This strand of protests literature is strongly linked to the information aggregation literature that builds on the modern versions of the celebrated Condorcet Jury Theorem, as formalized by Austen-Smith & Banks (1996), Feddersen & Pesendorfer (1997) and Myerson (1998b). For example, Battaglini (2017) considers a model where a biased incumbent similar to ours is trying to learn the state of the world (i.e. establish political communication) from the information that is dispersed among the citizens via an informal election, like a protest or a petition. They then pin down the conditions under which political communication is possible. Similarly, Ekmekci et al. (2019) study a model of protests where there maybe activists who derive extra utility just because they are participating in the protest, called 'activists'. The citizens add noise to the turnout and have implications for the information content of the protests.

3.3 The Base Model

There is a continuum of citizens with measure one, indexed by $j \in [0, 1]$, and an information provider, denoted by P. The citizens are considering joining a potential protest against an incumbent. The state of the world is denoted by θ and is distributed normally with mean μ_0 and standard deviation σ_0, i.e. $\theta \sim N(\mu_0, \sigma_0^2)$. This prior belief about the state of the world is shared by all players of the game. For simplicity, we normalize $\mu_0 = 0$ and $\sigma_0 = 1$. These normalizations do not qualitatively affect any of our results. The state of the world can be interpreted as any metric that pertains to the performance of the incumbent, where higher states imply poorer incumbent performance. It can also be interpreted as a general measure of anti-incumbent sentiment in the population.

Prior to their protest participation decision, each citizen receives a private signal from the information provider, denoted by $\theta_j \in \mathbb{R}$, where $\theta_j = \theta + \varepsilon_j$, $\varepsilon_j \sim N(0, \sigma_p^2)$ and ε_j's are independently and identically distributed. Hence, the private signal is equal to the true state of the world in expectation, even though each of them contain a random noise. The information provider determines $\sigma_p \in \mathbb{R}_{\geq 0} \cup \infty$, the standard deviation of the noise term in the signal. With a slight abuse of notation, we assume that if $\sigma_p = \infty$, then the information provider sends a random real number as a signal[1]. Hence, in effect, the information provider is 'truthful' in expectation, in the sense that the signal is expected to reveal the true state of the world, but it is free to choose how noisy the private signals are going to be. If the provider chooses $\sigma_p = 0$, then it reveals the true state of the world to all citizens, and similarly if it chooses $\sigma_p = \infty$, then the signal has no informational content. We assume that, as stated above, the information provider does not know the true state of the world and commits to a signal precision σ_p before the state of the world realizes. Notice that σ_p does not depend on the citizen identifier, i.e. whatever noise level the information provider commits to prior to the realization of the signals, it is going to affect all private signals in the same way.

After receiving their private signal θ_j, each citizen updates her prior belief about the state of the world θ according to the Bayes rule and takes an action $a_j \in \{0, 1\}$, where 0 denotes not joining the protest and 1 denotes joining the protest. Let n be the total measure of the citizens who decide to join the protest, i.e. $n = \int_0^1 a_j dj$. The protest succeeds if $n \geq t$ where $t \in [0, 1]$ is a predetermined and exogenous success threshold for the protest. We assume that this is common knowledge among the players of the game. We also assume that if the protest fails, the incumbent survives and if the protest is successful, the incumbent leaves office. Let $O \in \{S, F\}$ denote the outcome of the game, where S denotes the outcome of a successful protest and F denotes the outcome of a failed protest.

We assume that the information provider belongs to one of two types: a type that is trying to instigate the protest, denoted by P_i, and a type that is trying to dissuade the protest, denoted by P_d. We assume that sending noisy signals is costless for the information provider. To simplify the results, we normalize that P_i

[1]This asumption presents no technical difficulties, since these improper signal distribution with infinite mass are well-behaved in our case in the sense that they lead to proper posteriors given our proper prior, as discussed by Hartigan (2012). Also see the proof of Lemma 3.3.2 to see that the posterior distribution, which is proper in the limit.

gets utility 1 if the protest is successful and 0 otherwise. Similarly, P_d gets utility -1 if the protest is successful and 0 otherwise. These normalizations regarding the utility values for the information provider do not drive our results, but clearly costless noisy signals assumption does.

Let the payoffs for the citizens be denoted by $u_a^O(\theta_j)$. In other words, the signals affect the citizen's payoff from the protest participation decision along with affecting their beliefs about the state of the world. Furthermore, let $q(.)$ be the equilibrium probability of the protest succeeding. A representative citizen j who received the private signal θ_j is going to join the protest, i.e. $a_j^* = 1$ if and only if

$$q(.)u_1^S(\theta_j) + (1 - q(.))u_1^F(\theta_j) \geq q(.)u_0^S(\theta_j) + (1 - q(.))u_0^F(\theta_j)$$

where the left hand side of the inequality corresponds to the citizen's expected utility from joining the protest and the right hand side of the inequality corresponds to the citizen's expected utility from not joining the protest, given her private signal and the equilibrium probability of a successful protest. From this, we can pin down a lower bound for the equilibrium probability of the protest succeeding as a function of her private signal. Label this lower bound as $1 - h(\theta_j)$ where $h(\theta_j)$ is the probability of the regime surviving that makes the citizen who received private signal θ_j indifferent between joining the protest and not joining the protest:

$$q(.) \geq \frac{u_0^F(\theta_j) - u_1^F(\theta_j)}{(u_1^S(\theta_j) - u_0^S(\theta_j)) + (u_0^F(\theta_j) - u_1^F(\theta_j))} \equiv 1 - h(\theta_j)$$

For simplicity, we assume $u_1^S(\theta_j) = \theta_j$, $u_1^F(\theta_j) = \theta_j - 1$ and $u_0^F(\theta_j) = u_0^S(\theta_j) = 0$. These assumptions do not qualitatively affect the results of this paper. The results would be the same as long as *(i)* the relative value of participation is higher when the protest is successful, and this difference is increasing in the private signal value, *(ii)* the relative value of participation is increasing in the private signal value and *(iii)* the always exists some citizens who have the dominant strategy to join the protest and some citizens who have the dominant strategy to not join the protest. However, these assumptions lead us to the following simplifying equality: $h(\theta_j) = \theta_j$. This is straightforward from plugging in the assumed utility values to the inequality above. Notice that this also means that there are going to be two types of 'fanatic' citizens. One group of the citizens, who are the ones with relatively higher realizations of the signal, is going to derive a really high utility from participating in the protest, regardless of whether it is successful in the end or not. Similarly, another type of the citizens, who are the ones with relatively lower realizations of the signal, is going to derive a really high utility from not participating in the protest, regardless of whether it is successful in the end or not.

We restrict our attention to the symmetric Perfect Bayesian Equilibrium of the protest game defined above, as it is usually the case in the literature. With slight abuse of notation, let a protest participation strategy for citizen j be $a_j : \mathbb{R} \to [0, 1]$, which is a measurable function from her signal to her probability of joining the protest. Before moving to the equilibrium structure of the strategies of the citizens, we can make the following statements regarding the behavior of the citizens regardless of the value of the equilibrium dependent $q(.)$.

Lemma 3.3.1. *(i) A citizen who receives the private signal $\theta_j > 1$ has a dominant strategy to join the protest.*

(ii) A citizen who receives the private signal $\theta_j < 0$ has a dominant strategy not to join the protest.

(iii) For the citizens with $\theta_j \in [0, 1]$, the ones with higher private signal require a lower equilibrium probability of success to join the protest.

The lemma above is directly related to two-sided limit dominance assumption as presented in Morris & Shin (2001), and is an inherent feature of the global games models.

3.3.1 Equilibrium Strategies of the Citizens

Next, we turn to the equilibrium strategies for the citizens. The equilibrium strategy profile is going to have a cutoff structure where there exists a cutoff private signal $\widehat{\theta} \in [0,1]$ such that every citizen with a private signal higher than $\widehat{\theta}$ joins the protest, every citizen with a private signal lower than $\widehat{\theta}$ does not join the protest and the citizen with the private signal of exactly $\widehat{\theta}$ is indifferent between joining the protest and not joining the protest. The lemma below pins down this cutoff for the citizens equilibrium strategies.

Lemma 3.3.2. *There exists a unique $\widehat{\theta} \in [0,1]$ such that citizens with $\theta_j > \widehat{\theta}$ protest, and $\widehat{\theta}$ is pinned down by*

$$\Phi\left(\frac{\sigma_p}{\sqrt{1+\sigma_p^2}}\widehat{\theta} + \sqrt{1+\sigma_p^2}\Phi^{-1}(t)\right) = \widehat{\theta} \tag{3.1}$$

where Φ denotes the cumulative probability distribution of the standard normal distribution.

The lemma pins down the cutoff private signal with an equation which cannot be explicitly solved. The equation states that for an agent who receives the cutoff signal, the probability that the protest will be successful, the left hand side of the equation, is equal to the probability of successful protest that is required for the agent to be indifferent between joining and not joining, which is equal to her private signal. This equation in some form is prevalent in the global games literature. In sum, any given parameter t and action of the information provider σ_p induce a unique cutoff strategy $\widehat{\theta}$ for the citizens.

The idea of the proof is as follows. Given the Bayesian updating procedure of the citizens and the known distribution of the private signals, we can find the mass of citizens who join the protest. Using this, for any given cutoff strategy, we can back out a threshold state of the world where it leads to a successful protest. This, along with the distribution of the state of the world, give us the equilibrium probability that the protest is successful. Lastly, the cutoff value pinned down by

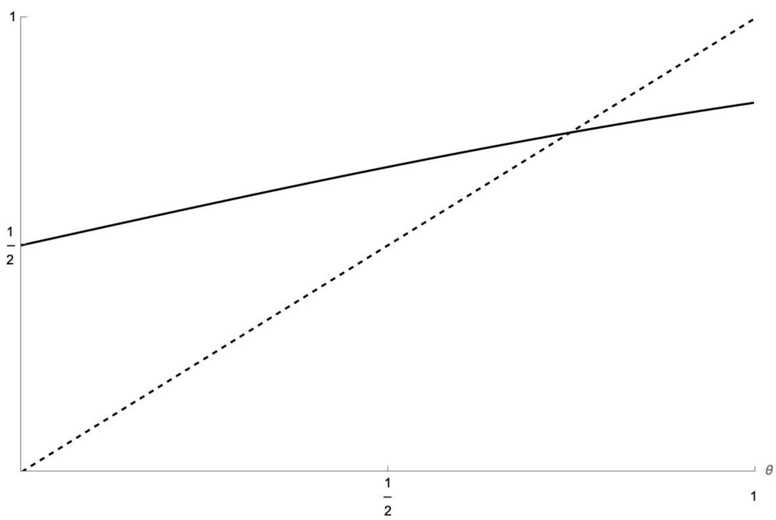

Figure 3.1: The figure plots the left hand side of the equation 3.1 (with the solid line) and the right hand side of the same equation (with the dashed line) when $\sigma_p = 2$ and $t = 0.5$. The intersection point of the two lines determine $\widehat{\theta}$.

the indifference condition of the citizen who receives this cutoff signal.

The lemma below presents some comparative statics regarding this equilibrium strategy with a cutoff signal for citizens, $\widehat{\theta}$, which is also equal to the probability of the regime surviving for the agent who received that cutoff signal.

Lemma 3.3.3. *(i) $\widehat{\theta}$ is increasing in σ_p if and only if $t > \Phi\left(-\frac{\widehat{\theta}}{(1+\sigma_p^2)\sigma_p}\right)$.*

(ii)

$$\lim_{\sigma_p \to \infty} \widehat{\theta} = \begin{cases} 1 & \text{if } t > \frac{1}{2} \\ t^f & \text{if } t = \frac{1}{2} \\ 0 & \text{if } t < \frac{1}{2} \end{cases}$$

where t^f is the fixed point of the standard normal cumulative distribution function and $\lim_{\sigma_p \to 0} \widehat{\theta} = t$.

(iii) $\widehat{\theta}$ is concave in σ_p if $t < \frac{1}{2}$ and strictly increasing in σ_p if $t \geq \frac{1}{2}$.

The lemma shows the change in the cutoff signal $\widehat{\theta}$, which is also equal to the probability of the protest being successful for the cutoff citizen, with respect to the noise parameter σ_p, and provides some limit values for the cutoff signal. It turns out that $\widehat{\theta}$ is increasing in σ_p if and only if t is high enough. To see why,

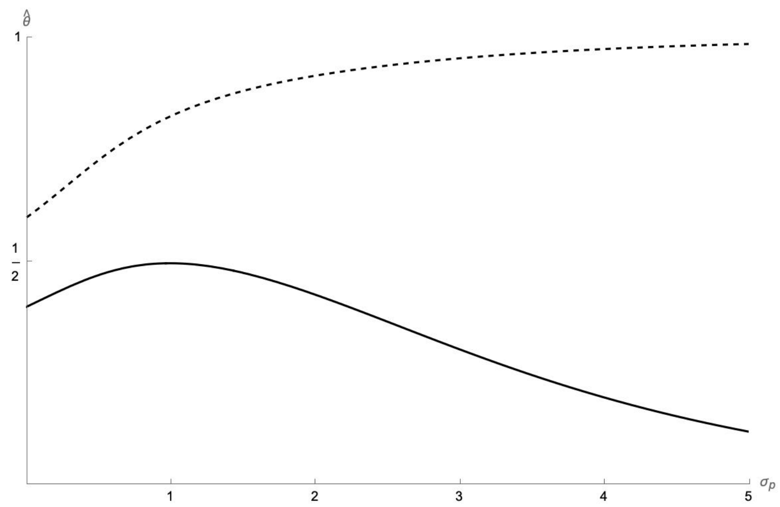

Figure 3.2: The figure plots $\widehat{\theta}$ as a function of σ_p for $t = 0.4$ (with the solid line) and for $t = 0.6$ (with the dashed line). The figure demonstrate the result in Lemma 3.3.3 part *(iii)*.

notice that an increase in σ_p leads to the following changes in the behavior of the citizens. First, a noisier signal is going to lead to a minimal update in the belief, since the new information is not reliable. In other words, the citizens are going to rely on their prior more. Remember that the prior leads them to expect that the state of the world is 0. Consider the agent with the cutoff signal $\widehat{\theta}$. After receiving the signal, her posterior belief is going to be $\frac{\widehat{\theta}}{1+\sigma_p}$. Notice that as σ_p gets larger, the citizen's posterior belief about the state of the world gets smaller and smaller, and converges to her prior 0 as σ_p goes to infinity. Since she was indifferent between protesting and not protesting before by definition, now she is going to have an incentive to refrain from to protest. This is going to lead to a rightward shift in the cutoff signal in equilibrium. Second, as a result of an increase in σ_p, the tail values of σ_p ('really high' and 'really low' signal realizations) have higher probability. Since $u_1^S(\theta_j) = \theta_j$, $u_1^F(\theta_j) = \theta_j - 1$ and $u_0^F(\theta_j) = u_0^S(\theta_j) = 0$, this means that effectively there is now a higher chance of 'fanatic' citizens who gain a lot of utility from participating in the protest and not participating in the protest. Combined together, if the required participation for a successful protest t is low enough, the 'fanatics' with very high signal realizations are going to be enough to create a successful protest, overcoming the negative effect coming from the less reliable signal. But they are not going to be enough if t is high enough.

3.3.2 Equilibrium Strategies of the Information Provider

Given this equilibrium play by the citizens, we now turn to the optimal strategy of the information provider. Given any σ_p picked by the provider and the true state of the world θ, the mass of citizens who protest is $n = \Phi\left(\frac{\theta-\hat{\theta}}{\sigma_p}\right)$. Then, given a cutoff strategy $\widehat{\theta}$, there exists a unique state of the world θ' such that the protest is successful if and only if $\theta \geq \theta'$, which is pinned down by $\Phi(\frac{\theta'-\hat{\theta}}{\sigma_p}) = t$ which implies $\theta' = \widehat{\theta} + \sigma_p \Phi^{-1}(t)$. Then, from the perspective of an information provider who is committing to a noise parameter according to the common prior belief, the probability of the protest being successful is $1 - \Phi\left(\widehat{\theta} + \sigma_p \Phi^{-1}(t)\right)$, which yields utility 1 to the P_i type information provider and -1 to the P_d type information provider, and the outcome of a failed protest yields the utility of 0. Given these, and fact that $\widehat{\theta}$ is determined by equation 3.1, the type P_d information provider's problem is

$$\max_{\sigma_{P_d}} -\left(1 - \Phi\left(\widehat{\theta} + \sigma_{P_d}\Phi^{-1}(t)\right)\right)$$

and similarly, the type P_i information provider's problem is

$$\max_{\sigma_{P_i}} 1 - \Phi\left(\widehat{\theta} + \sigma_{P_i}\Phi^{-1}(t)\right)$$

Given these observations, we provide the following result regarding the optimal signal noise by the P_d type information provider, who is trying to dissuade the protest.

Theorem 3.3.4. *For the P_d type information provider, there exists $\widehat{t} < \frac{1}{2}$, such that*

(i) if $t > \frac{1}{2}$, then $\sigma^*_{P_d} = \infty$. In equilibrium, $\widehat{\theta} = 1$ and the protest participation is half of the population, which means that the protest fails. The utility of the information provider is 0.

(ii) if $t = \frac{1}{2}$*, then* $\sigma^*_{P_d} = \infty$*. In equilibrium,* $\widehat{\theta} = t^f \approx 0.78$ *and the probability that the protest is successful is* $1 - t^f \approx 0.22$*. The expected utility of the information provider is* $t^f - 1 \approx -0.22$*.*

(iii) if $t \in (\widehat{t}, \frac{1}{2})$*, then* $\sigma^*_{P_d} \in (0, \infty)$*. In equilibrium,* $\widehat{\theta} \in (t, t^f)$ *and the probability that the protest is successful is strictly less than* $1 - \Phi(t)$*. The expected utility of the information provider is strictly higher than* $\Phi(t) - 1$*.*

(iv) if $t \leq \widehat{t}$*, then* $\sigma^*_{P_d} = 0$*. In equilibrium,* $\widehat{\theta} = t$ *and the probability that the protest is successful is* $1 - \Phi(t)$*. The expected utility of the information provider is* $\Phi(t) - 1$*.*

Theorem 3.3.4 reveals equilibrium behavior of a type P_d information provider, and the corresponding protest success probabilities. The P_d type information provider is able to completely dissuade the protest when the protest success threshold is higher than half of the population. It achieves this by sending signals that are completely uninformative. Citizens are left to rely on their priors, and since the protest success threshold is relatively high, the protest fails. The behavior of type P_d information provider is similar when the protest success threshold is exactly half of the population, the optimal signals are completely uninformative. But in that case, it is not able to dissuade the protest with probability 1. Instead, equation 3.1 reduces to $\Phi(\widehat{\theta}) = \widehat{\theta}$, i.e. the equilibrium cutoff signal is pinned down by the fixed point of the cumulative distribution function of the standard normal distribution, which is approximately equal to 0.78. This leads to a protest success probability of 0.22. Hence, if the required participation for a successful protest is greater than or equal to the half of total mass of citizens, then an information provider who is trying to dissuade the protest is going to send completely uninformative signals, forcing the citizens to rely on their priors for their protest participation decision. This is optimal because the citizen prior regarding θ is that it is distributed according to the standard normal distribution.

The equilibrium behavior of the P_d information provider changes when the protest success threshold is less than half. When t is low, the viewers are going to rely on their prior in case the signals are completely uninformative, which means that the protest is definitely going to be successful since we have the prior mean of 0. Hence, when t is low, the best that the P_d type provider can do is to commit to revealing the state of the world and hope that it is low enough to prevent a successful protest. But interestingly, when the threshold is between $\widehat{t}$ and half, where $\widehat{t} \approx 0.43$, the information provider now sends informative signals that do not completely reveal the state of the world. This is due to the concavity of the objective function of the P_d type information provider when the threshold is in that region. Neither fully revealing the state of the world nor sending completely uninformative signals is optimal. In fact, there is an optimal amount of obfuscation in terms of the informational quality of the signals that the provider can achieve to induce the lowest probability of a successful protest.

Next, we turn to the corresponding result for the P_i type information provider, who is trying to instigate the protest.

Theorem 3.3.5. *For the P_i type information provider,*

(i) if $t > \frac{1}{2}$, then $\sigma^*_{P_i} = 0$. In equilibrium, $\widehat{\theta} = t$ and the probability that the protest is successful is $1 - \Phi(t)$, which is also equal to the expected utility of the information provider.

(ii) if $t = \frac{1}{2}$, then $\sigma^*_{P_i} = 0$. In equilibrium, $\widehat{\theta} = \frac{1}{2}$ and the probability that the protest is successful is $1 - \Phi\left(\frac{1}{2}\right) \approx 0.31$, which is also equal to the expected utility of the information provider.

(iii) if $t < \frac{1}{2}$, then $\sigma^*_{P_i} = \infty$. In equilibrium, $\widehat{\theta} = 0$ and the protest participation is half the population, which means that the protest succeeds. The utility of the information provider is 1.

Similarly, Theorem 3.3.5 reveals equilibrium behavior of a type P_i information provider, and the corresponding protest success probabilities. The P_i type information provider is able to always instigate the protest when the protest success threshold is lower than half of the population. Similar to the P_d type information provider in the case of $t > \frac{1}{2}$, it achieves this by sending signals that are completely uninformative. As a result, citizens are left to rely on their priors, and since the protest success threshold is relatively low, the protest succeeds.

The equilibrium behavior of the P_i information provider changes when the protest success threshold is more than half. As opposed to sending completely uninformative signals in the case of less than half protest success threshold, the information provider now sends signals that completely reveal the state of the world. Notice that if the P_i type provider sends completely uninformative signals again, the citizens are going to rely on their priors, which means that the protest is never successful since the success threshold is more than half now. Hence that cannot be optimal. If it completely reveals the state of the world, then equation 3.1 reduces to $\Phi(t) = \widehat{\theta}$, which implies that the probability that the protest is successful is $1-\Phi(t)$, which is clearly an improvement over the previous case. But as revealed in Lemma 3.3.3, $\widehat{\theta}$ is strictly increasing in the noise parameter when $t \geq \frac{1}{2}$. Hence it is optimal for P_i type information provider to completely reveal the state of the world when $t \geq \frac{1}{2}$. This implies that the optimal strategy for the type P_i information provider when the success threshold is relatively high is essentially to commit to fully revealing the state of the world and hope for the best. Notice that, since the probability that the protest is successful is equal to $1-\Phi(t)$ when $t \geq \frac{1}{2}$, which is strictly decreasing over that parameter domain, and 1 otherwise, these results imply that the information provider is not able to set the noise parameter so that a failed protest is always more likely than a successful protest, regardless of the success threshold, due to the fact that $1 - \Phi\left(\frac{1}{2}\right) \approx 0.31$, which means that when $t \geq \frac{1}{2}$, a successful protest is less likely than a failed protest.

3.3.3 Citizen Welfare under Strategic Information Provider

Last in our exploration of the base model, we study the ex-ante welfare of the citizens given the equilibrium strategies. Notice that the ex-ante expected utility of a representative citizen, denoted by W, is given by

$$W = \int_{\hat{\theta}}^{\infty} \left(q(\theta_j;\hat{\theta})\theta_j + (1 - q(\theta_j;\hat{\theta}))(\theta_j - 1) \right) \phi(\theta_j) d\theta_j$$

where

$$q(\theta_j;\hat{\theta}) = 1 - \Phi\left(\frac{\sqrt{1+\sigma_p^2}}{\sigma_p}\widehat{\theta} + \sqrt{1+\sigma_p^2}\Phi^{-1}(t) - \frac{\theta_j}{\sqrt{1+\sigma_p^2}\sigma_p} \right)$$

We can simplify this expression, which is provided in the lemma below.

Lemma 3.3.6. *The ex-ante expected welfare of a citizen, denoted by W, can be written as*

$$W = \phi(\widehat{\theta}) - \int_{\widehat{\theta}}^{\infty} (1 - q(\theta_j;\widehat{\theta}))\phi(\theta_j)d\theta_j$$

Moreover, we have

$$\phi(\widehat{\theta}) \geq W > \phi(\widehat{\theta}) - \widehat{\theta}(1 - \Phi(\widehat{\theta}))$$

which means W is always positive.

Equipped with this simplified expression, we first pin down the value of W under the special cases of σ_p where it is equal to 0 and ∞.

Lemma 3.3.7. *The limit values of ex-ante citizen welfare can be expressed with*

$$\lim_{\sigma_p \to \infty} W = \begin{cases} \phi(1) - (1 - \Phi(1)) & \text{if } t > \frac{1}{2} \\ \phi(t^f) - t^f\left(1 - \Phi(t^f)\right) & \text{if } t = \frac{1}{2} \\ \phi(0) & \text{if } t < \frac{1}{2} \end{cases}$$

and

$$\lim_{\sigma_p \to 0} W = \phi(t)$$

where all of the values are positive. Moreover, $\lim_{\sigma_p \to 0} W > \lim_{\sigma_p \to \infty} W$ for $t \geq \frac{1}{2}$ and $\lim_{\sigma_p \to \infty} W > \lim_{\sigma_p \to 0} W$ for $t < \frac{1}{2}$.

The lemma above shows that the citizen welfare is positive both when the signals are uninformative and when they perfectly reveal the state. More importantly, the citizen welfare when the signals reveal the state is not always higher than the citizen welfare when the signals are uninformative. More specifically, if the participation threshold for a successful protest is higher than half, then signals that reveal the state are more beneficial for the citizen welfare, whereas if it is less than half, then signals that contain no information are more beneficial for the citizen welfare. To see why, notice that the citizens reach their highest possible utility in equilibrium when they participate in a successful protest. Since the protest is successful when the participation threshold for a successful protest is less than half and they rely on their priors, they prefer not to obtain any information as opposed to learning the true state of the world, since it can turn out to be too low for a successful protest even though the threshold is low. Conversely, when the protest is successful when the participation threshold for a successful protest is more than half, only thing that they can hope for a successful protest is that the underlying state of the world is high enough for enough people to join, which means learning the state of the world precisely is welfare enhancing compared to learning nothing about the state of the world.

Combining the two lemmas above, we reach the following corollary.

Corollary 3.3.7.1. *For a given t, the citizen welfare attains its highest value when $\sigma_p^* = 0$ and its lowest value when $\sigma_p^* = \infty$.*

The corollary above provides an important observation about ex-ante citizen welfare. It is at its highest when the information provider commits to fully revealing the state of the world, and it is at its lowest when the information provider commits to sending completely uninformative signals. This means that, for any intermediary value of $\sigma_p \in (0, \infty)$, the citizen welfare cannot be better than the case where $\sigma_p = 0$ or worse than the case ware $\sigma_p = \infty$, even though we do not know exactly how citizen welfare responds to marginal changes.

Combining this result with the Theorems 3.3.4 and 3.3.5, we reach the following result regarding the relationship between the type of the information provider and the citizen welfare.

Corollary 3.3.7.2. *P_i type information provider is more beneficial for citizen welfare in equilibrium compared to P_d type information provider if and only if $t < \frac{1}{2}$.*

This result is somewhat expected, since the citizens reach their highest possible utility when they participate in a successful protest. Since the P_i type information provider's objective is to facilitate a successful protest, the citizens derive the highest utility through P_i type information provider.

3.4 Extended Model with Strategic Incumbent

In this section, we turn to the second step of the objective of this paper, which is to explore political communication between an incumbent and the citizens when the information provider is strategic, as outlined in the preceding section. Hence, from here on, the incumbent is a strategic player as well.

We assume that there is an incumbent who benefits from remaining in office as long as the state of the world θ is not too large. Moreover, the incumbent does not know the true value of θ, but the citizens receive noisy private signals from an information provider, as explained in the preceding sections. Hence, in order to learn the state of the world, the incumbent can rely on the protest turnout, and stay in or leave office according to the outcome. Hence, in effect, the incumbent strategically determines a protest success threshold t and commits to resign if the protest is indeed successful. Notice that the incumbent is quasi-benevolent: she does not care purely her own power and she finds it beneficial to leave if the state of the world is actually too high. The main question we are investigating is the possibility of political communication: Can the incumbent learn the state of the world from the protests? And if yes, to what extent?

Formally, assume that the incumbent's utility from remaining in the office is defined by $U = b - \theta$ where $b > 0$ and her utility from leaving her office is 0.

We interpret b as a measure of bias between the incumbent and the citizens. We assume that the incumbent shares the prior of the information provider and the citizens, i.e. her prior regarding the state of the world is also $\theta \sim N(0,1)$. Given her utilities, the incumbent sets a protest success threshold $t \in [0,1]$, and she commits to leaving her office if and only if $n \geq t$. Given this value of t, the information provider sets σ_p, where the value depends on its type. Given t and σ_p, the signals of the citizens realize and they choose to join or refrain from the protest.

Notice that, if the incumbent knew the state of the world θ, she would leave her office if $\theta \geq b$, and remain in office otherwise. As stated above, she does not, hence she has to learn about the state of the world from the protest. But the information she can gather from the protest depends on both the noise parameter set by the information provider and the citizens' participation decision. Hence, she has to consider the effect of t on both of these. Given the equilibrium play of the information provider and the citizens, and the incumbent's choice of t, the probability of the protest being unsuccessful is $\Phi\left(\widehat{\theta} + \sigma_p \Phi^{-1}(t)\right)$, where $\widehat{\theta}$ is determined by equation 3.1 and σ_p is determined by Theorem 3.3.4 if the information provider is type P_d and Theorem 3.3.5 if the information provider is type P_i. Moreover, given her prior, the expected value of the state of the world when the protest is unsuccessful is $\int_{-\infty}^{\widehat{\theta}+\sigma_p\Phi^{-1}(t)} \theta_i \phi(\theta_i) d\theta_i = -\phi\left(\widehat{\theta} + \sigma_p \Phi^{-1}(t)\right)$[2]. Given these, the incumbent's problem is

$$\max_t \Phi\left(\widehat{\theta} + \sigma_p \Phi^{-1}(t)\right)\left(b + \phi\left(\widehat{\theta} + \sigma_p \Phi^{-1}(t)\right)\right)$$

where σ_p and $\widehat{\theta}$ are equilibrium objects.

Before pinning down the optimal action of the incumbent in the next subsection, we investigate the responses of the optimal actions by the citizens and the information provider to the incumbent's actions. First, we look at how $\widehat{\theta}$ changes with respect to the changes in t.

[2] Please refer to the proof of Lemma 3.3.6 to see why this equality holds.

Lemma 3.4.1. *(i)* $\widehat{\theta}$ *is increasing in* t.

(ii) $\lim_{t\to 1}\widehat{\theta} = 1$ *and* $\lim_{t\to 0}\widehat{\theta} = 0$.

The lemma shows the change in the cutoff signal $\widehat{\theta}$, which is also equal to the probability of the protest being successful for the cutoff citizen, with respect to the participation threshold for the protest success t, and also provides the limit values for the cutoff signal as t goes to its border values. It turns out that $\widehat{\theta}$ is increasing in t, which is expected. Since a higher t implies that it is less probable to participate in a successful protest, the agent with the cutoff signal would find not participating strictly better as a result of an increase in t.

Next, we turn to how σ_p changes with respect to the changes in t.

Lemma 3.4.2. *(i) For a type* P_d *provider,* $\sigma^*_{P_d}$ *is increasing in* t.

(ii) For a type P_i *provider,* $\sigma^*_{P_i}$ *is decreasing in* t.

The lemma above shows that, as a result of an increase in the threshold for a successful protest, a P_d type information provider provides less informative signals whereas a P_i type information provider provides more informative signals. To see why, notice that since a P_d type information provider is trying to dissuade the protest, if the threshold is really high, all it needs to do to achieve its goal is to send uninformative signals that forces them to rely on their priors, in which case they refrain from participating the protest. This is in contrast with the case of a P_i type information provider, who is trying to instigate a protest. Hence, as the threshold gets higher and higher, a P_i type information provider sends highly informative signals that reveal the state of the world, and hopes that it is high enough to lead to a successful protest.

3.4.1 Equilibrium Strategies of the Incumbent

Informed with these comparative statics results, we next turn to the problem of the incumbent. Given the optimal strategies of the information provider and the citizens, what is the level of protest success threshold that the incumbent needs to commit in order to maximize her utility? These optimal protest success thresholds, which depend on the level of incumbent bias as one might expect, are provided in the theorem below.

Theorem 3.4.3. *(i) When the information provider is P_d type, there exists $\underline{b} \approx 0.32$ and $\overline{b} \approx 1.06$ such that, for $b < \underline{b}$ we have a unique $t^* \in \left(\widehat{t}, \frac{1}{2}\right)$, for $b \in [\underline{b}, \overline{b}]$ we have $t^* = \frac{1}{2}$, and for $b > \overline{b}$ any $t > \frac{1}{2}$ is optimal.*

(ii) When the information provider is P_i type, there exists $\widetilde{b} \approx 0.6$ such that, if $b < \widetilde{b}$ there exists a unique $t^* > \frac{1}{2}$, and for $b \geq \widetilde{b}$ we have $t^* = 1$.

The theorem above reveals the optimal protest success threshold for an incumbent. If the information provider is trying to dissuade a protest, the incumbent knows that any threshold above simple majority leads to completely uninformative signals. Hence, as long as the bias is sufficiently low, the incumbent is going to try to learn from the protest by picking t^* less than half. If the bias is large, then the incumbent prefers to remain in office regardless, hence is going to choose some t^* greater than half, since any such t^* leads to a failed protest. Similarly, if the information provider is trying to instigate a protest, the incumbent knows that any threshold below simple majority leads to completely uninformative signals. Hence in equilibrium, she chooses a threshold that is greater than half. But as the bias grows, the optimal threshold grows as well, since the incumbent is less and less incentivized to leave office.

The theorem above leads to the following corollary about how the optimal threshold responds to changes in the incumbent bias, which directly follows from the proof of the theorem.

Corollary 3.4.3.1. *t^* is increasing in b.*

As expected, as the incumbent's bias grows, she is going to gain a higher utility by remaining in office, hence she is going to set a higher protest success threshold.

Given the optimal protest success threshold, the following corollary defines the equilibrium play and outcome of the game for given information provider type and bias. Again, since this result follows directly from Theorems 3.3.4, 3.3.5 and 3.4.3, it is provided without a proof.

Corollary 3.4.3.2. *(i) If the information provider is type P_d and (a) $b < \underline{b}$, then we have unique $t^* \in \left(\widehat{t}, \frac{1}{2}\right)$, $\sigma^*_{P_d} < \infty$, $\widehat{\theta} \in (t^*, t^f)$ and the incumbent remains in office with probability $\Phi(t^*)$, or (b) $b \in [\underline{b}, \overline{b}]$, then we have $t^* = \frac{1}{2}$, $\sigma^*_{P_d} = \infty$, $\widehat{\theta} = t^f$ and the incumbent remains in office with probability t^f or (c) $b > \overline{b}$, then we have any $t > \frac{1}{2}$ optimal, $\sigma^*_{P_d} = \infty$, $\widehat{\theta} = 1$ and the incumbent remains in office.*

(ii) If the information provider is type P_i, then we have unique $t^* > \frac{1}{2}$ when $b < \widetilde{b}$ and $t^* = 1$ when $b \geq \widetilde{b}$, $\sigma^*_{P_d} = 0$ regardless of the value of b, $\widehat{\theta} = t^*$ and the incumbent remains in office with probability $\Phi(t^*)$.

Interestingly, an information provider who is trying to instigate a protest always reveals the true state of the world in the equilibrium of the game, regardless of the value of b. To see why, notice that a P_i type information provider can directly create a successful protest by providing completely uninformative signals if the incumbent sets a threshold that is less than half. Since the incumbent knows this, she never sets such a threshold. However, when the incumbent sets a threshold higher than half, the P_i type information provider's best-case scenario is to commit to completely revealing the state of the world, and hope that it is high enough to create a successful protest. Hence, if the incumbent sets any threshold higher than half, she knows that the true state is going to be reflected in the protest outcome. Hence, her optimal threshold decision is directly determined by how biased she is, since the state is going to be revealed anyways, regardless of her choice.

However, the case with an information provider who is trying to dissuade a protest is much different. Given the main objective of the information provider, if the threshold is more than half, it sends completely uninformative signals, since in that case the protest is going to fail. Knowing this, a highly biased incumbent directly sets a success threshold that is higher than half, to ensure that she remains in office. If her bias is low, she faces the trade-off of learning about the state of the world (since a higher threshold leads to lower noise) and remaining in office in a state where it is not beneficial for her (since a higher threshold shrinks the space of states where she leaves office).

3.4.2 Political Communication

In this penultimate subsection, we turn to the motivating question of this section, which is the possibility of political communication. In simple terms, we are asking whether it is possible for the information provider to learn about the state of the world from the outcome of the game. In order to answer this question, we first define what constitutes a successful political communication.

Definition 3.4.4. *Political communication is possible if and only if the protest turnout reveals the state of the world, i.e.* $\sigma_p^* < \infty$.

Given any signal noise parameter committed by the information provider that that does not correspond to completely uninformative signals, it is possible to recover the state of the world just by looking at the cutoff signal, as explained in Section 3. This observation follows directly from the Law of Large Numbers. Hence, we say that political communication is possible if and only if the incumbent can recover the true state of the world by looking at the protest participation. Notice that political communication occurs if and only if the equilibrium is informative, since in equilibrium the citizens participate according to their signal, and given there are infinitely many citizens, the state of the world estimated by looking at the protest participation is arbitrarily precise (see, for an example, Battaglini (2017)).

It is important to note that political communication being possible does not imply that the incumbent can get the same utility as she does under perfect information when she knows the state of the world, since she is going to learn about the state of the world ex-post, hence her prior commitment regarding the protest success threshold may differ. We come back to this point at the end of the subsection.

Given this definition for political communication, we provide the following corollary, which follows directly from Theorem 3.4.3.

Corollary 3.4.4.1. *Political communication is possible if and only if either the information provider is P_i type or the information provider is P_d type and $b < \overline{b}$.*

Since the incumbent learns about the state of the world as long as the information provider sends signals that contain some information value, we can directly observe which parameters lead to political communication in equilibrium. Since the information provider who is trying to instigate a protest perfectly reveals the state of the world in equilibrium, the political communication always occurs under this type of provider. However, the information provider who is trying to dissuade the protest sends completely uninformative signals when the protest success threshold is larger than half, which in turn occurs in equilibrium if the incumbent's bias is large. Hence, under a P_d type information provider, political communication does not occur if the incumbent bias is larger than the $\underline{b}$ threshold. Remember that in this case, the incumbent's equilibrium behavior depends more on her bias than her desire to learn about the state of the world and make the right decision, since this case also leads to a failed protest with probability 1 (except the special case of the success threshold of half, where the probability of a failed protest is t^f). Hence, this leads to a failed political communication, since the incumbent is not necessarily interested in political communication to begin with, and the information provider's best interest is to obfuscate the state of the world to prevent a protest from happening. Notice that the same is happening in the case if a P_i type information provider as well, since the incumbent sets t^* equal to 1. But political communication is still possible since the information provider countervails this via committing to perfectly revealing the state of the world in order to instigate a successful protest.

Notice that all of these results assume that the incumbent commits to a protest success threshold and cannot renege on it after the fact. If it was the case that her decisions in the equilibrium of this game are the same as what they would have been in the hypothetical case where she knows the state of the world (hence, leaving office if and only if the state of the world is bigger than her bias), we could say that the assumption of commitment has no effect on the results, and if they differ, we can conclude that the assumption is vital for our results. Hence, in order to answer this question, we investigate the cases where the incumbent's decision under perfect information is the same as her decision under imperfect information.

In order to understand the effects of commitment on the use of information by the incumbent, we first define when she regrets her decisions ex-post, in the sense that she would have changed her action (leaving versus staying in office) if she knew the state.

Definition 3.4.5. *Given the state of the world, we say that the incumbent regrets her equilibrium action t^* if the result it leads to yields a lower payoff than the result she would have gotten under her equilibrium play with perfect information.*

Notice that this definition enables us to pin down the states of the world under which the incumbent would have an incentive to change her choice of protest success threshold ex-post, since she would do so if she regrets her action according to the definition above.

Equipped with this formulation, we present the following result that pins down the states where the incumbent regrets her decision, and the outcomes that lead to her regret.

Lemma 3.4.6. *t^* leads to the incumbent leaving office if and only if $\theta \geq b$, if and only if the information provider is type P_d and $b = t^f$. Furthermore, if*

(i) the information provider is type P_d and (a) $b < \underline{b}$ or $b \in (t^f, \overline{t}]$, then there exists states where the incumbent regrets leaving office, or (b) $b \in [\underline{b}, t^f)$ or $b > \overline{b}$, then there exists states where the incumbent regrets not leaving office.

(ii) the information provider is type P_i and (a) and $b < \widetilde{b}$ then there exists states where the incumbent regrets not leaving office, or (b) and $b \geq \widetilde{b}$ then there exists states where the incumbent regrets leaving office.

This result implies that, except a knife-edge case, the incumbent almost always disregards some information equilibrium, in the sense that her action does not match what her action would have been under perfect information. Notice that her bias has no effect on this observation, since she has the same bias under both cases. Hence, this observation is purely the result of the rule of commitment. Hence, since the incumbent has to commit to a protest success threshold ex-ante, she necessarily leaves some information content from the protest on the table.

As the lemma above also suggests, the assumption of commitment plays an important role in the results provided in this paper, since there is only one knife-edge case where the incumbent's decision under perfect information is the same as her decision under imperfect information, which only happens if the information provider is trying to dissuade the protest and her bias is equal to $t^f \approx 0.78$. In all other cases, the incumbent either regrets leaving office in some states, or not leaving office in some states.

3.4.3 Citizen Welfare under Strategic Incumbent

Lastly, we go back to investigating citizen welfare, now that we incorporated both a strategic incumbent and a strategic information provider. We study the ex-ante welfare of the citizens given the equilibrium strategies of them and the aforementioned agents. As presented in Lemma 3.3.6, the ex-ante expected utility of a citizen is given by

$$W = \phi(\widehat{\theta}) - \int_{\widehat{\theta}}^{\infty} (1 - q(\theta_j; \widehat{\theta}))\phi(\theta_j) d\theta_j$$

where

$$q(\theta_j; \hat{\theta}) = 1 - \Phi\left(\frac{\sqrt{1+\sigma_p^{*^2}}}{\sigma_p^*}\widehat{\theta} + \sqrt{1+\sigma_p^{*^2}}\Phi^{-1}(t^*) - \frac{\theta_j}{\sqrt{1+\sigma_p^{*^2}}\sigma_p^*}\right)$$

where $\widehat{\theta}$ is given by equation 3.1, σ_p^* is given by Theorems 3.3.4 and 3.3.5, depending on the information provider type, and t^* is given by Theorem 3.4.3.

Our focal point in this subsection is the effect of the incumbent bias on the ex-ante citizen welfare, which is pinned down by the lemma below.

Lemma 3.4.7. *The ex-ante welfare of a citizen is decreasing in b.*

As expected, the citizen welfare is decreasing in the incumbent bias. This is straightforward since, as the bias increases, there is a bigger mass of states where the citizens would prefer if the incumbent left office, but the incumbent is not willing to do so.

3.5 Conclusion

In this paper, we first set up a global games model of protests where the information provider is trying to influence the result of the protest, where the result is success if the turnout is higher than the success threshold determined by the incumbent and failure otherwise. We find that, if the information provider is trying to dissuade the protest, it is going to send completely uninformative signals if the protest success threshold is more than half, and send partially or fully informative signals if it is less than half. This is in contrast to the case of an information provider who is trying to instigate a protest, since it sends signals that completely reveal the state of the world when the protest success threshold is more than half, and sends completely uninformative signals if it is less than half. Hence, as expected, different types of information providers who have opposite aims behave in opposite ways.

Next, we turn the incumbent into a strategic player as well, and we assume that she chooses the protest success threshold to maximize her utility that depends on the state of the world, but also her bias level. We find that the incumbent's equilibrium action has a cutoff structure in her bias under both types of information providers. More specifically, when the information provider is trying to dissuade the protest, if the incumbent bias is low, then she sets a success threshold that is less than half, if the bias is at a medium level, then she sets the success threshold at exactly half, and if the bias is relatively large, then any threshold that is higher than half is optimal since they all lead to her staying in office with probability 1. When the information provider is trying to instigate a protest, if the incumbent bias is relatively low, she sets a success threshold higher than half, and if the bias is relatively high, then she requires that all citizens attend the protest to deem it successful.

Lastly we look at when political communication is possible, in the sense that when it is possible for the incumbent to learn the state of the world by looking at the protest outcome. We find that the political communication is possible except when the information provider is trying to dissuade the protest and the incumbent bias is high. We also show that her actions almost never match what they would be under complete information. We reach this result by showing that, except in a knife-edge case, there are always states of the world where she regrets the outcome, in the sense that she either regrets leaving office or staying in office. We interpret this result as the incumbent disregarding some of the informational content of the protest due to the fact that she has to commit to the protest success threshold before the protest occurs. So even though the state of the world is revealed, she does not fully incorporate it to her actions.

For future work, one of the most important extensions would be regarding the commitment assumptions we have, both for the incumbent and the information provider, since the last result suggests that it plays a vital role in driving our results. Furthermore, a more general formulation of the citizen utilities and priors would be beneficial to reach more general results. Even though none of our results qualitatively depends on them, they could still reveal interesting observations. This also applies to the incumbent's utility, since we assume a very simple objective function for the incumbent. It is an open question how the incumbent would behave if she had different objectives.

Appendix

Proof of Lemma 3.3.1. *(i)* Given that $h(\theta_j) = \theta_j$, a citizen with the private signal θ_j joins the protest if and only if $q(.) \geq 1 - \theta_j$. Given that $q(.)$ is a probability between 0 and 1, $\theta_j \geq 1 - q(.)$ always holds for $\theta_j > 1$. Hence it is a dominant strategy for the citizen to join the protest.

(ii) Similarly, $\theta_j \geq 1 - q(.)$ never holds for $\theta_j < 0$. Hence it is a dominant strategy for the citizen not to join the protest.

(iii) For citizens with $\theta_j \in [0, 1]$, who do not have a dominant strategy, their strategies depend on equilibrium object $q(.)$. The lowest value of the required equilibrium probability of the protest succeeding is pinned down by $1 - \theta_j$ for an agent with private signal θ_j, which is decreasing in θ_j. The desired result follows. ■

Proof of Lemma 3.3.2. Notice that, given our assumptions regarding the utilities of the citizens, a citizen who received the private signal θ_j is going to join the protest if and only if $q(.)\theta_j + (1 - q(.))(\theta_j - 1) \geq 0$, which holds if and only if $\theta_j - (1 - q(.)) > 0$. Hence, the equilibrium strategy for the citizens has a cutoff structure where there exists a cutoff private signal $\widehat{\theta} \in [0, 1]$ such that every citizen with a private signal higher than $\widehat{\theta}$ joins the protest, every citizen with a private signal lower than $\widehat{\theta}$ does not join the protest and the citizen with the private signal of exactly $\widehat{\theta}$ is indifferent between joining the protest and not joining the protest. This indifference condition is pinned down by $q(\theta_j = \widehat{\theta}, \widehat{\theta}) = 1 - \widehat{\theta}$.

Take a citizen who received a private signal θ_j, and denote her posterior belief regarding θ by $f(\theta|\theta_j)$. After observing the signal and updating her beliefs, her posterior will be distributed normally with mean $\frac{\theta_j}{1+\sigma_p^2}$ and standard deviation $\frac{\sigma_p}{\sqrt{1+\sigma_p^2}}$, i.e.

$$f(\theta|\theta_j) = \frac{1}{\frac{\sigma_p}{\sqrt{1+\sigma_p^2}}} \phi \left(\frac{\theta - \frac{\theta_j}{1+\sigma_p^2}}{\frac{\sigma_p}{\sqrt{1+\sigma_p^2}}} \right)$$

where ϕ is the probability density function for the standard normal distribution.

Moreover, since the private signals are normally distributed with mean 0 and standard deviation σ_p, the probability that a randomly selected citizen joins the protest, which is also equal to the mass of citizens who join the protest, is pinned down by $\Phi\left(\frac{\theta-\hat{\theta}}{\sigma_p}\right)$. Then, given a cutoff strategy $\widehat{\theta}$ as explained above, there exists a unique state of the world θ' such that the protest is successful if and only if $\theta \geq \theta'$, which is pinned down by $\Phi(\frac{\theta'-\hat{\theta}}{\sigma_p}) = t$ which implies $\theta' = \widehat{\theta} + \sigma_p \Phi^{-1}(t)$. Given this, the citizen's belief about the protest being successful is

$$P(\theta \geq \theta'|\theta_j) = q(\theta_j, \widehat{\theta}) = 1 - \Phi\left(\frac{\widehat{\theta} + \sigma_p \Phi^{-1}(t) - \frac{\theta_j}{1+\sigma_p^2}}{\frac{\sigma_p}{\sqrt{1+\sigma_p^2}}}\right)$$

Then the indifference condition, pinned down by $q(\theta_j = \widehat{\theta}, \widehat{\theta}) = 1 - \widehat{\theta}$, can be rewritten as

$$\Phi\left(\frac{\sigma_p}{\sqrt{1+\sigma_p^2}}\widehat{\theta} + \sqrt{1+\sigma_p^2}\Phi^{-1}(t)\right) = \widehat{\theta}$$

Notice that since a cumulative distribution function can take on values only between 0 and 1, this means that $\widehat{\theta} \in [0, 1]$.

Now we need to show that there indeed exists and equilibrium and it is unique. Define

$$g(\widehat{\theta}) = \widehat{\theta} - \Phi\left(\frac{\sigma_p}{\sqrt{1+\sigma_p^2}}\widehat{\theta} + \sqrt{1+\sigma_p^2}\Phi^{-1}(t)\right)$$

Notice that $g(\widehat{\theta})$ is continuous in $\widehat{\theta}$, strictly increasing, $\lim_{\widehat{\theta}\to-\infty} g(\widehat{\theta}) = -\infty$ and $\lim_{\widehat{\theta}\to\infty} g(\widehat{\theta}) = \infty$. These together imply that there exists a unique $\widehat{\theta}$ such that $g(\widehat{\theta}) = 0$, which is the same $\widehat{\theta}$ that satisfies the indifference condition above. ■

Proof of Lemma 3.3.3. (i) Totally differentiating equation 3.1, we get

$$\phi\left(\frac{\sigma_p}{\sqrt{1+\sigma_p^2}}\widehat{\theta} + \sqrt{1+\sigma_p^2}\Phi^{-1}(t)\right)\left(\frac{\widehat{\theta}}{(1+\sigma_p^2)^{\frac{3}{2}}} + \frac{\sigma_p}{\sqrt{1+\sigma_p^2}}\frac{\partial\widehat{\theta}}{\partial\sigma_p} + \frac{\sigma_p}{\sqrt{1+\sigma_p^2}}\Phi^{-1}(t)\right)$$
$$= \frac{\partial\widehat{\theta}}{\partial\sigma_p}$$

Rearranging, we get

$$\phi\left(\frac{\sigma_p}{\sqrt{1+\sigma_p^2}}\widehat{\theta}+\sqrt{1+\sigma_p^2}\Phi^{-1}(t)\right)\left(\frac{\widehat{\theta}}{(1+\sigma_p^2)^{\frac{3}{2}}}+\frac{\sigma_p}{\sqrt{1+\sigma_p^2}}\Phi^{-1}(t)\right)$$
$$=\left(1-\frac{\phi\left(\frac{\sigma_p}{\sqrt{1+\sigma_p^2}}\widehat{\theta}+\sqrt{1+\sigma_p^2}\Phi^{-1}(t)\right)\sigma_p}{\sqrt{1+\sigma_p^2}}\right)\frac{\partial\widehat{\theta}}{\partial\sigma_p}$$

Hence we have

$$\frac{\partial\widehat{\theta}}{\partial\sigma_p}=\frac{\phi\left(\frac{\sigma_p}{\sqrt{1+\sigma_p^2}}\widehat{\theta}+\sqrt{1+\sigma_p^2}\Phi^{-1}(t)\right)\left(\frac{\widehat{\theta}}{(1+\sigma_p^2)^{\frac{3}{2}}}+\frac{\sigma_p}{\sqrt{1+\sigma_p^2}}\Phi^{-1}(t)\right)}{\left(1-\frac{\phi\left(\frac{\sigma_p}{\sqrt{1+\sigma_p^2}}\widehat{\theta}+\sqrt{1+\sigma_p^2}\Phi^{-1}(t)\right)\sigma_p}{\sqrt{1+\sigma_p^2}}\right)}$$

Notice that the expression in the denominator is positive since $\frac{\sigma_p}{\sqrt{1+\sigma_p^2}}$ is bounded above by 1. The numerator is positive if and only if $\frac{\widehat{\theta}}{(1+\sigma_p^2)^{\frac{3}{2}}}+\frac{\sigma_p}{\sqrt{1+\sigma_p^2}}\Phi^{-1}(t)>0$, which holds if and only if $t>\Phi\left(-\frac{\widehat{\theta}}{(1+\sigma_p^2)\sigma_p}\right)$.

(ii) Notice that $\lim_{\sigma_p\to\infty}\frac{\sigma_p}{\sqrt{1+\sigma_p^2}}\widehat{\theta}=\widehat{\theta}$ and

$$\lim_{\sigma_p\to\infty}\sqrt{1+\sigma_p^2}\Phi^{-1}(t)=\begin{cases}\infty & \text{if } t>\frac{1}{2}\\ 0 & \text{if } t=\frac{1}{2}\\ -\infty & \text{if } t<\frac{1}{2}\end{cases}$$

we have

$$\lim_{\sigma_p\to\infty}\widehat{\theta}=\begin{cases}1 & \text{if } t>\frac{1}{2}\\ t^f & \text{if } t=\frac{1}{2}\\ 0 & \text{if } t<\frac{1}{2}\end{cases}$$

where t^f is the fixed point of the standard normal cumulative distribution function, approximately equal to 0.78. Similarly, we have $\lim_{\sigma_p\to 0}\frac{\sigma_p}{\sqrt{1+\sigma_p^2}}\widehat{\theta}=0$ and

$\lim_{\sigma_p \to 0} \sqrt{1+\sigma_p^2}\Phi^{-1}(t) = \Phi^{-1}(t)$ we have $\lim_{\sigma_p \to 0} \widehat{\theta} = t$.

(iii) Notice that $\lim_{\sigma_p \to 0} \Phi\left(-\frac{\widehat{\theta}}{(1+\sigma_p^2)\sigma_p}\right) = 0$. This means that $\lim_{\sigma_p \to 0} \frac{\partial \widehat{\theta}}{\partial \sigma_p} > 0$. This means that $\widehat{\theta}$ is always increasing in σ_p when σ_p is close to 0. Moreover $\lim_{\sigma_p \to \infty} \Phi\left(-\frac{\widehat{\theta}}{(1+\sigma_p^2)\sigma_p}\right) = \frac{1}{2}$. This means that as σ_p diverges to infinity, $\widehat{\theta}$ is still increasing in σ_p if and only if $t \geq \frac{1}{2}$. Since $\Phi\left(-\frac{\widehat{\theta}}{(1+\sigma_p^2)\sigma_p}\right)$ is monotonic in σ_p, the desired result follows. ■

Proof of Theorem 3.3.4. (i) If $t \geq \frac{1}{2}$, then $\widehat{\theta}$ is strictly increasing in σ_p. This means that the P_d type information provider's objective function, $-\left(1 - \Phi\left(\widehat{\theta} + \sigma_p \Phi^{-1}(t)\right)\right)$ is strictly increasing in σ_p, which leads to $\sigma_{P_d}^* = \infty$. Plugging this into equation 3.1, we can find that $\widehat{\theta} = 1$. Moreover, since the protest is successful if and only if the state of the world is greater than $\widehat{\theta} + \sigma_p \Phi^{-1}(t)$, we can conclude that the protest is never successful. Moreover, plugging in $\sigma_p = \infty$ into the expected utility of the type P_d information provider, we can find that it is 0.

(ii) In the special case of $t = \frac{1}{2}$, equation 3.1 boils down to

$$\Phi\left(\frac{\sigma_p}{\sqrt{1+\sigma_p^2}}\widehat{\theta}\right) = \widehat{\theta}$$

and the probability that the protest is successful is $1 - \Phi(\widehat{\theta})$. This means that the expected utility of the information provider is $-(1 - \Phi(\widehat{\theta}))$, which is increasing in $\widehat{\theta}$. Since $\frac{\sigma_p}{\sqrt{1+\sigma_p^2}}$ is increasing in σ_p and $\lim_{\sigma_p \to \infty} \sqrt{1+\sigma_p^2} = 1$, the maximum value of $\widehat{\theta}$ is pinned down by $\Phi(\widehat{\theta}) = \widehat{\theta}$, i.e. the fixed point of the cumulative distribution function of the standard normal distribution, which is approximately 0.78. Plugging this value into the expressions above yields the stated values in the theorem.

(iii) If $t < \frac{1}{2}$, then $\widehat{\theta}$ is concave in σ_p, which implies $-\left(1 - \Phi\left(\widehat{\theta} + \sigma_p \Phi^{-1}(t)\right)\right)$ is quasi-concave in σ_p since Φ is strictly increasing. As a result, $\sigma_p^* > 0$ if and only if the derivative of the objective function of the P_d type information provider is positive as the limit as $\sigma_p \to 0$, since then we can say that $\sigma_p = 0$ is dominated by some

positive σ_p. Few simple steps show that $\lim_{\sigma_p \to 0} \frac{\partial}{\partial \sigma_p} - \left(1 - \Phi\left(\widehat{\theta} + \sigma_p \Phi^{-1}(t)\right)\right) = \phi\left(\Phi^{-1}(t)\right) t + \Phi^{-1}(t)$. Hence, there exists a $\widehat{t} \in (0, \frac{1}{2})$ that is pinned down by the condition $\phi\left(\Phi^{-1}(t)\right) t + \Phi^{-1}(t) = 0$ such that $\sigma_p^* > 0$ if and only if $t \in (\widehat{t}, \frac{1}{2})$. Since $\sigma_p^* > 0$, through Lemma 3.3.3 and the fact that $\widehat{\theta} = t$ when $\sigma_p = 0$ and $\widehat{\theta} = t^f$ when $t = \frac{1}{2}$, we can conclude that $\widehat{\theta} \in (t, t^f)$ in equilibrium. These together imply that the success probability $1 - \Phi(\widehat{\theta} + \sigma_p \Phi^{-1}(t))$ is less than $1 - \Phi(t)$. Plugging that in, we can find that the expected utility of the type P_d information provider is bounded below by $\Phi(t) - 1$.

(iv) If $t < \underline{t}$, then the objective function is strictly decreasing in σ_p due to reasons explained in the previous part, hence we have $\sigma_p^* = 0$. This implies $\widehat{\theta} = t$. These together imply that the success probabilityis $1 - \Phi(t)$. Plugging that in, we can find that the expected utility of the type P_d information provider is $\Phi(t) - 1$ ■

Proof of Theorem 3.3.5. (i) If $t > \frac{1}{2}$, then $\widehat{\theta}$ is strictly increasing in σ_p. This means that the P_i type information provider's objective function, $1 - \Phi\left(\widehat{\theta} + \sigma_p \Phi^{-1}(t)\right)$ is strictly decreasing in σ_p, which leads to $\sigma_{P_d}^* = 0$. Plugging that into equation 3.1, we get $\widehat{\theta} = t$. Plugging this into the relevant equations above, we get the expressions in the theorem.

(ii) In the special case of $t = \frac{1}{2}$, equation 3.1 boils down to

$$\Phi\left(\frac{\sigma_p}{\sqrt{1 + \sigma_p^2}} \widehat{\theta}\right) = \widehat{\theta}$$

and the probability that the protest is successful is $1 - \Phi(\widehat{\theta})$, which is also equal to the expected utility of the information provider. Notice that this is decreasing in $\widehat{\theta}$. Since $\frac{\sigma_p}{\sqrt{1+\sigma_p^2}}$ is increasing in σ_p, we get $\sigma_p^* = 0$ and $\widehat{\theta} = \Phi(0) = \frac{1}{2}$. The probability that the protest is successful is $1 - \Phi(0.5) \approx 0.31$, which is also equal to the expected utility of the information provider.

(iii) If $t < \frac{1}{2}$, then $\widehat{\theta}$ is concave in σ_p, which implies $1 - \Phi\left(\widehat{\theta} + \sigma_p \Phi^{-1}(t)\right)$ is quasi-convex since $\sigma_p \Phi^{-1}(t)$ is linear in σ_p and cumulative distribution function is strictly increasing. This then means that one of the border solutions is optimal.

Notice that $\lim_{\sigma_p \to 0} 1 - \Phi\left(\widehat{\theta} + \sigma_p \Phi^{-1}(t)\right) = 1 - \Phi(t)$ since $\lim_{\sigma_p \to 0} \widehat{\theta} = t$. Moreover, $\lim_{\sigma_p \to \infty} 1 - \Phi\left(\widehat{\theta} + \sigma_p \Phi^{-1}(t)\right) = 1$ for $t < \frac{1}{2}$. Since $t < 1$, $\sigma_p = \infty$ yields a higher utility. Plugging that into equation 3.1, we get $\widehat{\theta} = 0$. This means that the protest participation is half the population since $\theta \sim N(0,1)$, which directly implies the protest is successful since $t < \frac{1}{2}$. ■

Proof of Lemma 3.3.6. First, notice that W simplifies to

$$
\begin{aligned}
W &= \int_{\hat{\theta}}^{\infty} \left(q(\theta_j; \hat{\theta}) + \theta_j - 1\right) \phi(\theta_j) d\theta_j \\
&= \int_{\hat{\theta}}^{\infty} q(\theta_j; \hat{\theta}) \phi(\theta_j) d\theta_j + \int_{\hat{\theta}}^{\infty} \theta_j \phi(\theta_j) d\theta_j - \int_{\hat{\theta}}^{\infty} \phi(\theta_j) d\theta_j \\
&= \int_{\hat{\theta}}^{\infty} q(\theta_j; \hat{\theta}) \phi(\theta_j) d\theta_j + \int_{\hat{\theta}}^{\infty} \theta_j \phi(\theta_j) d\theta_j - \left(1 - \Phi(\widehat{\theta})\right)
\end{aligned}
$$

Moreover, notice that

$$
\begin{aligned}
\int_{\hat{\theta}}^{\infty} \theta_j \phi(\theta_j) d\theta_j &= \int_{\hat{\theta}}^{\infty} \theta_j \frac{1}{\sqrt{2\pi}} e^{-\frac{\theta_j^2}{2}} d\theta_j \\
&= \frac{1}{\sqrt{2\pi}} \int_{\hat{\theta}}^{\infty} \theta_j e^{-\frac{\theta_j^2}{2}} d\theta_j \\
&= -\frac{1}{\sqrt{2\pi}} \int_{\hat{\theta}}^{\infty} e^{-\frac{\theta_j^2}{2}} d\left(-\frac{\theta_j^2}{2}\right) \quad \text{(since } -\theta_j d\theta_j = -d\left(-\frac{\theta_j^2}{2}\right)\text{)} \\
&= -\frac{1}{\sqrt{2\pi}} e^{-\frac{\theta_j^2}{2}} \Bigg|_{\widehat{\theta}}^{\infty} \\
&= \frac{1}{\sqrt{2\pi}} e^{-\frac{\widehat{\theta}^2}{2}} \\
&= \phi(\widehat{\theta})
\end{aligned}
$$

and

$$\int_{\hat{\theta}}^{\infty} q(\theta_j;\hat{\theta})\phi(\theta_j)d\theta_j = \int_{\hat{\theta}}^{\infty}\left(1-\Phi\left(\frac{\sqrt{1+\sigma_p^2}}{\sigma_p}\widehat{\theta}+\sqrt{1+\sigma_p^2}\Phi^{-1}(t)-\frac{\theta_j}{\sqrt{1+\sigma_p^2}\sigma_p}\right)\right)$$

$$\phi(\theta_j)d\theta_j$$

$$= (1-\Phi(\widehat{\theta}))$$

$$-\int_{\hat{\theta}}^{\infty}\Phi\left(\frac{\sqrt{1+\sigma_p^2}}{\sigma_p}\widehat{\theta}+\sqrt{1+\sigma_p^2}\Phi^{-1}(t)-\frac{\theta_j}{\sqrt{1+\sigma_p^2}\sigma_p}\right)\phi(\theta_j)d\theta_j$$

$$= (1-\Phi(\widehat{\theta}))-\int_{\hat{\theta}}^{\infty}\left(1-q(\theta_j;\widehat{\theta})\right)\phi(\theta_j)d\theta_j$$

Combining everything together, we have

$$\begin{aligned} W &= \int_{\hat{\theta}}^{\infty} q(\theta_j;\hat{\theta})\phi(\theta_j)d\theta_j + \int_{\hat{\theta}}^{\infty}\theta_j\phi(\theta_j)d\theta_j - \left(1-\Phi(\widehat{\theta})\right) \\ &= (1-\Phi(\widehat{\theta})) - \int_{\hat{\theta}}^{\infty}\left(1-q(\theta_j;\widehat{\theta})\right)\phi(\theta_j)d\theta_j + \phi(\widehat{\theta}) - \left(1-\Phi(\widehat{\theta})\right) \\ &= \phi(\widehat{\theta}) - \int_{\hat{\theta}}^{\infty}\left(1-q(\theta_j;\widehat{\theta})\right)\phi(\theta_j)d\theta_j \end{aligned}$$

The existence of the upper bound follows directly from the fact that $\int_{\hat{\theta}}^{\infty}\left(1-q(\theta_j;\widehat{\theta})\right)\phi(\theta_j)d\theta_j$ is positive valued for any value of $\widehat{\theta}$. To see the existence of the lower bound for W, notice that

$$1-q(\theta_j;\widehat{\theta}) < \Phi\left(\frac{\sigma_p}{\sqrt{1+\sigma_p^2}}\widehat{\theta}+\sqrt{1+\sigma_p^2}\Phi^{-1}(t)\right) = \widehat{\theta}$$

for $\theta_j > \widehat{\theta}$, following from the expression for $q(\theta_j;\widehat{\theta})$ and equation 3.1 and the fact that $1-q(\theta_j;\widehat{\theta})$ is decreasing in θ_j. Hence we have

$$\phi(\widehat{\theta}) - \int_{\hat{\theta}}^{\infty}\left(1-q(\theta_j;\widehat{\theta})\right)\phi(\theta_j)d\theta_j > \phi(\widehat{\theta}) - \widehat{\theta}\int_{\hat{\theta}}^{\infty}\phi(\theta_j)d\theta_j = \phi(\widehat{\theta}) - \widehat{\theta}(1-\Phi(\widehat{\theta}))$$

which yields the desired result. ■

Proof of Lemma 3.3.7. First, consider the value of W as $\sigma_p \to \infty$. Notice that

$$\lim_{\sigma_p \to \infty} \phi(\widehat{\theta}) = \begin{cases} \phi(1) & \text{if } t > \frac{1}{2} \\ \phi(t^f) & \text{if } t = \frac{1}{2} \\ \phi(0) & \text{if } t < \frac{1}{2} \end{cases}$$

which follows directly from Lemma 3.3.3. Moreover, we have

$$\lim_{\sigma_p \to \infty} 1 - q(\theta_j; \widehat{\theta}) = \begin{cases} 1 & \text{if } t > \frac{1}{2} \\ \Phi(t^f) = t^f & \text{if } t = \frac{1}{2} \\ 0 & \text{if } t < \frac{1}{2} \end{cases}$$

Combining all of the above and using Lemma 3.3.3 yield the first expression in the lemma.

Next, consider the value of W as $\sigma_p \to 0$. Notice that $\lim_{\sigma_p \to 0} \phi(\widehat{\theta}) = \phi(t)$ which follows directly from Lemma 3.3.3. Moreover, we have

$$\lim_{\sigma_p \to 0} 1 - q(\theta_j; \widehat{\theta}) = \begin{cases} 1 & \text{if } \widehat{\theta} > \theta_j \\ 0 & \text{otherwise} \end{cases}$$

where only the second line is relevant since the integral in the expression for W runs from $\widehat{\theta}$ to ∞. Combining everything implies $\lim_{\sigma_p \to 0} W = \phi(t)$.

Lastly, plugging in the values in the first expression shows that it is always below $\phi(t)$ for $t \geq \frac{1}{2}$. Moreover notice that $\phi(t)$ is decreasing in t, hence $\phi(0) > \phi(t)$ for all t, yielding the desired last result in the lemma. ■

Proof of Lemma 3.4.1. (i) Totally differentiating equation 3.1, we get

$$\phi\left(\frac{\sigma_p}{\sqrt{1+\sigma_p^2}}\widehat{\theta} + \sqrt{1+\sigma_p^2}\Phi^{-1}(t)\right)\left(\frac{\sigma_p}{\sqrt{1+\sigma_p^2}}\frac{\partial\widehat{\theta}}{\partial t} + \frac{\sqrt{1+\sigma_p^2}}{\phi(\Phi^{-1}(t))}\right) = \frac{\partial\widehat{\theta}}{\partial t}$$

Rearranging, we get

$$\phi\left(\frac{\sigma_p}{\sqrt{1+\sigma_p^2}}\widehat{\theta}+\sqrt{1+\sigma_p^2}\Phi^{-1}(t)\right)\frac{\sqrt{1+\sigma_p^2}}{\phi(\Phi^{-1}(t))}=$$
$$\left(1-\frac{\phi\left(\frac{\sigma_p}{\sqrt{1+\sigma_p^2}}\widehat{\theta}+\sqrt{1+\sigma_p^2}\Phi^{-1}(t)\right)\sigma_p}{\sqrt{1+\sigma_p^2}}\right)\frac{\partial\widehat{\theta}}{\partial t}$$

With the same arguments as part *(i)* of Lemma 3.3.3, we can see that $\frac{\partial\widehat{\theta}}{\partial t}>0$ always holds.

(ii) Notice that since $\lim_{t\to 0}\sqrt{1+\sigma_p^2}\Phi^{-1}(t) = -\infty$ and $\lim_{t\to 1}\sqrt{1+\sigma_p^2}\Phi^{-1}(t) = \infty$, we have $\lim_{t\to 1}\widehat{\theta} = 1$ and $\lim_{t\to 0}\widehat{\theta} = 0$ from equation 3.1. ■

Proof of Lemma 3.4.2. (i) Notice that $\sigma^*_{p_d} = \infty$ for $t \geq \frac{1}{2}$ and $\sigma^*_{p_d} \in (0,\infty)$ for $t < \frac{1}{2}$ as pinned down by Theorem 3.3.4. Hence we need to prove that $\sigma^*_{p_d}$ for $t \in (\widehat{t}, \frac{1}{2})$ is increasing in t. Remember that $\sigma^*_{p_d}$ for $t < \frac{1}{2}$ is pinned down by the first order condition

$$\phi\left(\widehat{\theta}+\sigma_p\Phi^{-1}(t)\right)\left(\frac{\partial\widehat{\theta}}{\partial\sigma_p}+\Phi^{-1}(t)\right)=0$$

which implies $\frac{\partial\widehat{\theta}}{\partial\sigma_p} = -\Phi^{-1}(t)$. Notice that as t is increasing, $\frac{\partial\widehat{\theta}}{\partial\sigma_p}$ is decreasing. Since $\widehat{\theta}$ is concave in σ_p when $t < \frac{1}{2}$, this means that $\sigma^*_{P_d}$ is increasing in t, since for a concave function, a lower derivative implies higher input value, which implies as t increases $\sigma^*_{P_d}$ must be increasing as well.

(ii) The result follows directly from the proof of Theorem 3.3.5. ■

Proof of Theorem 3.4.3. (i) Assume that the information provider is P_d type. First, notice that setting $t^* > \frac{1}{2}$ would yield utility b to the incumbent, setting $t^* = \frac{1}{2}$ would yield utility $t^f(b+\phi(t^f))$ and setting $t < \widehat{t}$ would yield $t(b+\phi(t))$ which then would lead to $t = \widehat{t}$ since that expression is increasing in t. But then notice that $t^f(b+\phi(t^f)) > \widehat{t}(b+\phi(\widehat{t}))$ for any b, hence setting $t \leq \underline{t}$ is strictly dominated. Moreover, setting $t^* < \frac{1}{2}$ would turn the problem into

$$\max_t \Phi\left(\widehat{\theta}+\sigma^*_{P_d}\Phi^{-1}(t)\right)\left(b+\phi\left(\widehat{\theta}+\sigma^*_{P_d}\Phi^{-1}(t)\right)\right)$$

where $\sigma^*_{P_d}$ is given by Theorem 3.3.4 part *(iii)* and $\widehat{\theta}+\sigma^*_{P_d}\Phi^{-1}(t) < t^f$ since $\Phi^{-1}(t) < 0$, $\sigma^*_{P_d} > 0$ and $\widehat{\theta} < t^f$ when $t < \frac{1}{2}$. In this case, the derivative of the objective function of the incumbent with respect to t is

$$\Big[\phi\left(\widehat{\theta}+\sigma^*_{P_d}\Phi^{-1}(t)\right)b+\phi\left(\widehat{\theta}+\sigma^*_{P_d}\Phi^{-1}(t)\right)^2 \\ +\Phi\left(\widehat{\theta}+\sigma^*_{P_d}\Phi^{-1}(t)\right)\phi'\left(\widehat{\theta}+\sigma^*_{P_d}\Phi^{-1}(t)\right)\Big]\left(\frac{\partial\widehat{\theta}}{\partial t}+\frac{\sigma^*_{P_d}}{\phi\left(\Phi^{-1}(t)\right)}\right)$$

where the term in the second parentheses is always positive, given the result of Lemma 3.4.1. Moreover, notice that the term is monotonic in b. This means that t^* is defined by

$$\phi\left(\widehat{\theta}+\sigma^*_{P_d}\Phi^{-1}(t^*)\right)b+\phi\left(\widehat{\theta}+\sigma^*_{P_d}\Phi^{-1}(t^*)\right)^2 \\ +\Phi\left(\widehat{\theta}+\sigma^*_{P_d}\Phi^{-1}(t^*)\right)\phi'\left(\widehat{\theta}+\sigma^*_{P_d}\Phi^{-1}(t^*)\right)=0$$

as long as there exists t^* for any given value of $b>0$ and $\widehat{\theta}+\sigma^*_{P_d}\Phi^{-1}(t^*) < t^f$. Moreover, since $\frac{\phi'\left(\widehat{\theta}+\sigma^*_{P_d}\Phi^{-1}(t^*)\right)}{\phi\left(\widehat{\theta}+\sigma^*_{P_d}\Phi^{-1}(t^*)\right)} = -\widehat{\theta}-\sigma^*_{P_d}\Phi^{-1}(t^*)$, we get

$$b=\Phi\left(\widehat{\theta}+\sigma^*_{P_d}\Phi^{-1}(t^*)\right)\left(\widehat{\theta}+\sigma^*_{P_d}\Phi^{-1}(t^*)\right)-\phi\left(\widehat{\theta}+\sigma^*_{P_d}\Phi^{-1}(t^*)\right)$$

Notice that the right hand side is takes the value 0 at $\widehat{\theta}+\sigma^*_{P_d}\Phi^{-1}(t^*) \equiv \widetilde{x} \approx 0.5 < t^f$, and strictly increasing thereon. This means that there exists $\underline{b} > 0$ such that for $b < \underline{b}$, t^* is pinned down by the equation above. At $\underline{b}$, the maximum value function is equal to $t^f(b+\phi(t^f))$. This implies that $\underline{b}$ is pinned down by $\underline{b} = {t^f}^2 - \phi(t^f) \approx 0.32$. Moreover, notice that $t^f(b+\phi(t^f)) > b$ if and only if

$b < \frac{t^f \phi(t^f)}{1-t^f} \equiv \overline{b} \approx 1.06$. This means that for $b \in (\underline{b}, \overline{b})$, $t = \frac{1}{2}$ is optimal, and for $b > \overline{b}$, any $t > \frac{1}{2}$ is optimal, since they all lead to $\sigma^*_{P_d} = \infty$ and $\widehat{\theta} = 1$.

(ii) Assume that the information provider is P_i type. This means that, given $\sigma^*_{P_i}$ as pinned down by Theorem 3.3.5, her objective function boils down to $\Phi(t)\,(b + \phi(t))$ when she chooses $t^* \geq \frac{1}{2}$ and 0 when she chooses $t^* < \frac{1}{2}$. First notice that $\Phi(t)\,(b + \phi(t)) > 0$ for any value of t and b, hence any $t^* < \frac{1}{2}$ is strictly dominated. Moreover, for t^* is pinned down by the first order condition $\phi(t^*)b + \phi(t^*)^2 + \Phi(t^*)\phi'(t^*) = 0$, which boils down to $t^* = \frac{b+\phi(t^*)}{\Phi(t)}$, given that $t^* \in \left[\frac{1}{2}, 1\right]$. Notice that at $b = 0$, we have $t^* > \frac{1}{2}$ and t^* is strictly increasing in b. Hence there exists a $\widetilde{b} = \Phi(1) - \phi(1) \approx 0.6$ such that for $b < \widetilde{b}$, there exists a unique t^* pinned down by $t^* = \frac{b+\phi(t^*)}{\Phi(t^*)}$ and for $b > \widetilde{b}$ we have $t^* = 1$. ■

Proof of Lemma 3.4.6. As explained above, if the incumbent knew the state of the world, she would leave her office if and only if $\theta \geq b$. In the equilibrium of the incomplete information game presented in this paper, she leaves office if and only if $\theta \geq \widehat{\theta} + \sigma^*_p \Phi^{-1}$. This means that the outcomes would coincide if and only if $b = \widehat{\theta} + \sigma^*_p \Phi^{-1}$.

First, assume that the information provider is type P_d. Furthermore, assume that $b < \underline{b}$. This means that, in equilibrium, the following condition holds.

$$b = \Phi\left(\widehat{\theta} + \sigma^*_{P_d}\Phi^{-1}(t^*)\right)\left(\widehat{\theta} + \sigma^*_{P_d}\Phi^{-1}(t^*)\right) - \phi\left(\widehat{\theta} + \sigma^*_{P_d}\Phi^{-1}(t^*)\right)$$

But this means that $b > \widehat{\theta} + \sigma^*_{P_d}\Phi^{-1}(t^*)$, which means that there are states where the incumbent regrets leaving office. Next, assume $b \in [\underline{b}, \overline{b}]$. This implies $\widehat{\theta} + \sigma^*_{P_d}\Phi^{-1}(t^*) = t^f$. Hence, the outcomes coincide if and only if $b = t^f$, which is possible since $t^f \in [\underline{b}, \overline{b}]$. If $b < t^f$, then there are states where the incumbent regrets not leaving office, and if $b > t^f$, then there are states where the incumbent regrets leaving office. Next, assume $b > \overline{b}$. This means that the incumbent never leaves office regardless of b, hence the outcomes are different, and there are states where she regrets not leaving office.

Next, assume that the information provider is type P_i. Then we have $\widehat{\theta} + \sigma^*_{P_i}\Phi^{-1}(t^*) = t^*$ where t^* is pinned down by $t^* = \frac{b+\phi(t^*)}{\Phi(t^*)}$ for $b < \widetilde{b}$, which implies $t^* > b$. Furthermore, for $b \geq \widetilde{b}$, we have $t^* = 1$, but $\widetilde{b} > 1$, hence we have $b > t^*$. This means that the outcomes never coincide, furthermore, there are states where the incumbent regrets not leaving office if $b < \widetilde{b}$ and there are states where the incumbent regrets leaving office if $b \geq \widetilde{b}$. ■

Proof of Lemma 3.4.7. As presented in Corollary 3.3.7.1, the ex-ante citizen welfare attains its highest value when $\sigma_p = 0$ and its lowest value when $\sigma_p = \infty$.

First assume that the information provider is type P_d. As presented in Corollary 3.3.7.1, the ex-ante citizen welfare attains its highest value when $\sigma_p = 0$ and its lowest value when $\sigma_p = \infty$. This means that the citizen welfare is higher when $b < \underline{b}$ where $\sigma^*_{P_d} < \infty$, as opposed to $b \geq \underline{b}$ where $\sigma^*_{P_d} = \infty$. Furthermore, when $\sigma^*_{P_d} = \infty$, we have, as presented in Lemma 3.3.7, $W = \phi(t^f) - t^f(1 - \Phi(t^f))$ if $b \in [\underline{b}, \overline{b}]$ and $W = \phi(1) - (1 - \Phi(1))$ when $b > \overline{b}$. The fact that $\phi(t^f) - t^f(1 - \Phi(t^f)) > \phi(1) - (1 - \Phi(1))$ concludes the proof for the P_d information provider type.

Now assume that the information provider is type P_i. Since $\sigma^*_{P_i} = 0$ in any equilibrium, we have, as presented in Lemma 3.3.7, $W = \phi(t^*)$. Since $t* > \frac{1}{2}$ when $b < \widetilde{b}$, $t* = 1$ when $b > \widetilde{b}$, and ϕ is strictly decreasing for $t^* \geq 0$, the proof follows. ■

References

Amenta, E., Elliott, T. A., Shortt, N., Tierney, A. C., Türkoğlu, D., & Vann Jr, B. (2017). From bias to coverage: What explains how news organizations treat social movements. *Sociology Compass*, *11*(3), e12460.

Anderson, D. L. (2002). *The columbia guide to the vietnam war.* Columbia University Press.

Arrow, K. J., Blackwell, D., & Girshick, M. A. (1949). Bayes and minimax solutions of sequential decision problems. *Econometrica, Journal of the Econometric Society*, 213–244.

Austen-Smith, D., & Banks, J. S. (1996). Information aggregation, rationality, and the condorcet jury theorem. *American political science review*, *90*(1), 34–45.

Battaglini, M. (2017). Public protests and policy making. *The Quarterly Journal of Economics*, *132*(1), 485–549.

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

Boix, C., & Svolik, M. W. (2013). The foundations of limited authoritarian government: Institutions, commitment, and power-sharing in dictatorships. *The Journal of Politics*, *75*(2), 300–316.

Bouton, L. (2013). A theory of strategic voting in runoff elections. *American Economic Review*, *103*(4), 1248–88.

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

Bouton, L., & Gratton, G. (2015). Majority runoff elections: strategic voting and duverger's hypothesis. *Theoretical Economics*, *10*(2), 283–314.

Brown, D. K., & Mourão, R. R. (2021). Protest coverage matters: How media framing and visual communication affects support for black civil rights protests. *Mass Communication and Society*, *24*(4), 576–596.

Burke, J. (2008). Primetime spin: Media bias and belief confirming information. *Journal of Economics & Management Strategy*, *17*(3), 633–665.

Carlsson, H., & Van Damme, E. (1993). Global games and equilibrium selection. *Econometrica: Journal of the Econometric Society*, 989–1018.

Chan, J., Lizzeri, A., Suen, W., & Yariv, L. (2018). Deliberating collective decisions. *The Review of Economic Studies*, *85*(2), 929–963.

Che, Y.-K., & Mierendorff, K. (2019). Optimal dynamic allocation of attention. *American Economic Review*, *109*(8), 2993–3029.

Chemerinsky, E., & Fisk, C. (1997). The filibuster. *Stanford L. Rev.*, *49*, 181.

Condorcet, M. d. (1785). Essay on the application of analysis to the probability of majority decisions. *Paris: Imprimerie Royale*.

Cooper, A. H. (2002). Media framing and social movement mobilization: German peace protest against inf missiles, the gulf war, and nato peace enforcement in bosnia. *European Journal of Political Research*, *41*(1), 37–80.

D'Alessio, D., & Allen, M. (2000). Media bias in presidential elections: A meta-analysis. *Journal of communication*, *50*(4), 133–156.

De Mesquita, E. B. (2010). Regime change and revolutionary entrepreneurs. *American Political Science Review*, *104*(3), 446–466.

Dewan, T., & Myatt, D. P. (2008). The qualities of leadership: Direction, communication, and obfuscation. *American Political science review*, *102*(3), 351–368.

Dewan, T., & Myatt, D. P. (2012). On the rhetorical strategies of leaders: Speaking clearly, standing back, and stepping down. *Journal of Theoretical Politics*, *24*(4), 431–460.

Duggan, J., & Martinelli, C. (2001). A bayesian model of voting in juries. *Games and Economic Behavior*, *37*(2), 259–294.

Edmond, C. (2013). Information manipulation, coordination, and regime change. *Review of Economic studies*, *80*(4), 1422–1458.

Ekmekci, M. (2009). Manipulation through political endorsements. *Journal of Economic Theory*, *144*(3), 1227–1248.

Ekmekci, M., Lauermann, S., et al. (2019). *Informal elections with dispersed information* (Tech. Rep.). University of Bonn and University of Mannheim, Germany.

Feddersen, T., & Pesendorfer, W. (1997). Voting behavior and information aggregation in elections with private information. *Econometrica: Journal of the Econometric Society*, 1029–1058.

Feddersen, T., & Pesendorfer, W. (1999). Elections, information aggregation, and strategic voting. *Proceedings of the National Academy of Sciences*, *96*(19), 10572–10574.

Fudenberg, D., Strack, P., & Strzalecki, T. (2018). Speed, accuracy, and the optimal timing of choices. *American Economic Review*, *108*(12), 3651–84.

Galperti, S., & Trevino, I. (2020). Coordination motives and competition for attention in information markets. *Journal of Economic Theory*, *188*, 105039.

Gentzkow, M., Shapiro, J. M., & Stone, D. F. (2015). Media bias in the marketplace: Theory. In *Handbook of media economics* (Vol. 1, pp. 623–645). Elsevier.

Gitlin, T. (2003). *The whole world is watching: Mass media in the making and unmaking of the new left.* Univ of California Press.

Goodin, R. E., & List, C. (2006). Special majorities rationalized. *British Journal of Political Science*, *36*(2), 213–241.

Groeling, T. (2013). Media bias by the numbers: Challenges and opportunities in the empirical study of partisan news. *Annual Review of Political Science*, *16*, 129–151.

Hartigan, J. A. (2012). *Bayes theory.* Springer Science & Business Media.

Hein, M. (2020). Do constitutional entrenchment clauses matter? constitutional review of constitutional amendments in europe. *International Journal of Constitutional Law*, *18*(1), 78–110.

Henry, E., & Ottaviani, M. (2019). Research and the approval process: The organization of persuasion. *American Economic Review*, *109*(3), 911–55.

Ke, T. T., & Villas-Boas, J. M. (2019). Optimal learning before choice. *Journal of Economic Theory*, *180*, 383–437.

Kim, S. H. (1993). "we (the supermajority of) the people": The development of a rationale for written higher law in north american constitutions. *Proceedings of the American Philosophical Society*, *137*(3), 364–389. Retrieved 2022-05-10, from http://www.jstor.org/stable/986998

Lipnowski, E., Mathevet, L., & Wei, D. (2020). Attention management. *American Economic Review: Insights*, *2*(1), 17–32.

Little, A. T., et al. (2012). Elections, fraud, and election monitoring in the shadow of revolution. *Quarterly Journal of Political Science*, *7*(3), 249–283.

Little, A. T., Tucker, J. A., & LaGatta, T. (2015). Elections, protest, and alternation of power. *The Journal of Politics*, *77*(4), 1142–1156.

Madestam, A., Shoag, D., Veuger, S., & Yanagizawa-Drott, D. (2013). Do political protests matter? evidence from the tea party movement. *The Quarterly Journal of Economics*, *128*(4), 1633–1685.

Mandler, M. (2012). The fragility of information aggregation in large elections. *Games and Economic Behavior*, *74*(1), 257–268.

Maug, E., & Yilmaz, B. (2002). Two-class voting: a mechanism for conflict resolution. *American Economic Review*, *92*(5), 1448–1471.

Mayskaya, T. (2022). Dynamic choice of information sources. *California Institute of Technology Social Science Working Paper, ICEF Working Paper WP9/2019/05*.

McAdam, D., & Su, Y. (2002). The war at home: Antiwar protests and congressional voting, 1965 to 1973. *American sociological review*, 696–721.

Morris, S., & Shin, H. S. (1998). Unique equilibrium in a model of self-fulfilling currency attacks. *American Economic Review*, 587–597.

Morris, S., & Shin, H. S. (2001). Global games: Theory and applications.

Morris, S., & Shin, H. S. (2002). Social value of public information. *american economic review*, *92*(5), 1521–1534.

Moscarini, G., & Smith, L. (2001). The optimal level of experimentation. *Econometrica*, *69*(6), 1629–1644.

Myatt, D. P. (2007). On the theory of strategic voting. *The Review of Economic Studies*, *74*(1), 255–281.

Myatt, D. P. (2017). A theory of protest voting. *The Economic Journal*, *127*(603), 1527–1567.

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

Myerson, R. B. (1998a). Extended poisson games and the condorcet jury theorem. *Games and Economic Behavior*, *25*(1), 111–131.

Myerson, R. B. (1998b). Population uncertainty and poisson games. *International Journal of Game Theory*, *27*(3), 375–392.

Nelson, T. E., Oxley, Z. M., & Clawson, R. A. (1997). Toward a psychology of framing effects. *Political behavior*, *19*(3), 221–246.

Nikandrova, A., & Pancs, R. (2018). Dynamic project selection. *Theoretical Economics*, *13*(1), 115–143.

Persson, T., & Tabellini, G. (2009). Democratic capital: The nexus of political and economic change. *American Economic Journal: Macroeconomics*, *1*(2), 88–126.

Prat, A. (2015). Media capture and media power. In *Handbook of media economics* (Vol. 1, pp. 669–686). Elsevier.

Prat, A., & Strömberg, D. (2013). The political economy of mass media. *Advances in economics and econometrics*, *2*, 135.

Puglisi, R., & Snyder Jr, J. M. (2015). Empirical studies of media bias. In *Handbook of media economics* (Vol. 1, pp. 647–667). Elsevier.

Schwartzberg, M. (2010). The arbitrariness of supermajority rules. *Social science information*, *49*(1), 61–82.

Schwartzberg, M. (2013). *Counting the many: The origins and limits of supermajority rule* (Vol. 10). Cambridge University Press.

Shadmehr, M., & Bernhardt, D. (2011). Collective action with uncertain payoffs: coordination, public signals, and punishment dilemmas. *American Political Science Review*, *105*(4), 829–851.

Sims, C. A. (2003). Implications of rational inattention. *Journal of monetary Economics*, *50*(3), 665–690.

Wald, A. (1947). Foundations of a general theory of sequential decision functions. *Econometrica, Journal of the Econometric Society*, 279–313.

Walgrave, S., & Vliegenthart, R. (2012). The complex agenda-setting power of protest: Demonstrations, media, parliament, government, and legislation in belgium, 1993-2000. *Mobilization: An International Quarterly*, *17*(2), 129–156.

Wit, J. (1998). Rational choice and the condorcet jury theorem. *Games and Economic Behavior*, *22*(2), 364–376.