Essays on Voting, Learning, and Dynamic Games

Inauguraldissertation

zur Erlangung des Grades eines Doktors
der Wirtschaftswissenschaften

durch

die Rechts- und Staatswissenschaftliche Fakultät der
Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Kailin Chen

aus Zherong, China

2024

Dekan:	Prof. Dr. Jürgen von Hagen
Erstreferent:	Prof. Dr. Stephan Lauermann
Zweitreferent:	Prof. Dr. Sven Rady
Tag der mündlichen Prüfung:	11. Juni 2024

Acknowledgements

This dissertation would not have been possible without the support and inspiration of a number of wonderful individuals. I appreciate all of them for being part of this journey.

First and foremost, I am deeply indebted to Stephan Lauermann, who was always extremely welcoming and generous with his time. Before meeting him, I had excelled at exams, but I had no experience in conducting research. He helped me start my career, inspired me to study economic theory, spent a lot of time and patience improving my writing, encouraged me when I was frustrated, and gave me the strength and confidence to continue. I would not have come this far without his constant support and unwavering guidance. When I have the opportunity to supervise students, I hope to be as supportive and caring as Stephan was to me.

I am exceptionally lucky to have Sven Rady as my supervisor. He was consistently kind to me, yet very strict and meticulous when it came to reviewing my writing. In addition, his assistance during my job search was exceptionally valuable. His support has been instrumental in shaping my academic journey. He is a role model to me as a researcher.

I would also like to extend my sincere gratitude to Benny Moldovanu and Johannes Hörner, who encouraged me to challenge existing results and provided me with valuable advice. I consider myself fortunate to have been a member of the amazing theory group at the Institute of Microeconomics, University of Bonn, which has been a continual source of encouragement and advice. I am especially grateful to Justus Preusser, Simon Block, Sophie Kreutzkamp, Carl Heese, Günnur Ege Bilgin, and Axel Niemeyer. They have been critical to my development as a scholar, and each has greatly helped me to refine my dissertation and presentation skills. I would also like to thank Sarah Lane, who provided numerous pieces of advice and help during my job search.

Finally, I am forever indebted to my parents for giving me the opportunities and experiences that have made me who I am. They have always supported me unconditionally to pursue my goals in academia. I am deeply grateful to my girlfriend and best companion, Huiyu Liu, who is always there for me, encouraging me during the most difficult time, and making a foreign city home.

Kailin Chen
Helsinki, Sep 2023

Contents

Acknowledgements iii

List of Figures vii

Introduction 1

1 Learning from Biased Souces 3

1.1 Introduction 3

1.2 Model 7

1.2.1 Basic Setting 7

1.2.2 Three-State Scenario 8

1.2.3 Strategy and Equilibrium 10

1.2.4 Characterization of Informative Equilibria 11

1.3 Information Aggregation 13

1.3.1 Results from the Existing Literature 13

1.3.2 Failure of Information Aggregation 15

1.4 Information Transmission 19

1.4.1 Existence of Informative Equilibria 19

1.4.2 Ranking Informative Equilibria 23

1.4.3 Amount of Information Transmission 24

1.4.4 Information Aggregation and Transmission 26

1.5 Beyond the Binary Situation 28

1.6 Commitment Case 33

1.7 Related Literature 36

1.8 Concluding Remarks 38

Appendix 1.A Proofs 38

1.A.1 Proof of Lemma 1.1, Proposition 1.2, Proposition 1.3 39

1.A.2 Informative Equilibrium with $\hat{T} = 1$ 42

1.A.3 Proof for the Results in Section 1.4 48

1.A.4 Mixed-Strategy Equilibria 49

1.A.5 Commitment Case 51

References 54

2 **Information Aggregation in Collective Experimentation** **57**

2.1 Introduction 57

2.2 Model 61

2.2.1 Model Setting 61

2.2.2 Equilibrium 62

2.2.3 Beliefs 62

2.3 Characterization 63

2.4 Large Elections 64

2.4.1 Limit Cut-off 64

2.4.2 Information Aggregation 65

2.5 Concluding Remarks 69

Appendix 2.A Proofs 69

References 73

3 **Fishing for Approval** **75**

3.1 Introduction 75

3.2 Model 76

3.2.1 Model Setting 76

3.2.2 Strategy 78

3.2.3 Preliminary Analysis 78

3.3 Equilibrium Analysis 79

3.3.1 Characterization 79

3.3.2 Large Elections 81

3.4 Observable Order 83

3.4.1 Symmetric Equilibrium 83

3.4.2 Asymmetric Equilibrium 85

3.5 Conclusion 87

Appendix 3.A Proofs 87

3.A.1 Proof of Proposition 3.1 88

3.A.2 Proof of Proposition 3.4 89

3.A.3 Set of Symmetric BNE 92

3.A.4 Set of Responsive Equilibria 93

References 97

List of Figures

1.1	Preferences	8
1.2	Preferences when the realized state is uncertain	9
1.3	Different thresholds of acceptance when $q_2 = 0$.	9
1.4	Inference from being pivotal and thresholds of acceptance	14
1.5	The distributions of the total number of approvals	16
1.6	The distributions of the total number of approvals	17
1.7	Information transmission and aggregation	22
1.8	Non-monotonic boundaries	23
1.9	Maximal amount of information transmission	26
1.10	Information transmission with the failure of information aggregation	27
1.11	The most likely pivotal frequency	33
1.12	Distributions of the total number of signal h	34
1.13	Distributions of the total number of signal h	36
2.1	Limit cut-off	64
2.2	Limit ratios of votes	67

Introduction

This dissertation collects three essays on microeconomic theory.

Chapter 1 studies learning from multiple informed agents where each agent has a small piece of information about the unknown state of the world in the form of a noisy signal and sends a message to the principal, who then makes a decision that is not constrained by predetermined rules. In contrast to the existing literature, I model the conflict of interest between the principal and the agents more generally and consider the case where the preferences of the principal and the agents are misaligned in some realized states. I show that if the conflict of interest between the principal and the agents is moderate, there is a discontinuity: when the number of agents is large enough, adding even a tiny probability of misaligned states leads to complete unraveling in which the agents ignore their signals, in contrast to the almost complete revealing that is predicted by the existing literature. Furthermore, I demonstrate that no matter how small the conflict of interest between the principal and the agents is, the information contained in each agent's message must vanish as the number of agents grows large. Finally, no matter how many agents there are, the total amount of information that is transmitted is limited, and the principal always fails to fully learn the unknown state.

Implementing a reform has divergent effects on a population and generates dispersed information concerning its overall suitability. Chapter 2 analyzes a collective experimentation model in which voters gradually learn their payoffs, which are divergent among them. Furthermore, their payoffs depend on the unknown state of the world. Hence, experimentation generates information concerning the unknown state, which is dispersed among the voters. I am interested in how strategic voting shapes incentives for experimentation, and more importantly, whether elections can aggregate and utilize the voters' private information concerning the unknown state. I show that a stricter rule for experimentation leads to more experimentation when the number of voters is large, and demonstrate information is effectively aggregated only if the voting rule is biased toward experimentation.

Chapter 3 is a joint work with Mehmet Ekmekci and Stephan Lauermann. We study a situation in which one success can make up for a long record of failure, and hence an individual might engage in a costly search to fish for one approval from decentralized approval agencies. We frame this situation as a sequential unanimity voting model since one approval agency's decision matters only if the others all reject the individual. The existing literature shows that the unanimity rule produces inefficiency in simultaneous and sequential elections. We further address the additional concern that the individual might fish for approval. Surprisingly, we show that the individual's ability to fish for approval helps the approval agencies to elicit information from him/her, leading to the approval agencies' first-best, full information outcome as the number of them grows large.

Chapter 1

Learning from Biased Souces

1.1 Introduction

This paper studies learning from multiple informed agents where each agent has a small piece of information about the unknown state of the world in the form of a noisy signal and sends a message to the principal, who then makes a decision that is not constrained by predetermined rules. This framework applies to scenarios that include non-binding shareholder voting, public protests, and survey polls in which the principal corresponds to the manager, politician, or interviewer, and the agents correspond to shareholders, citizens, or interviewees.

If the principal and the agents share the same preferences, then the agents report their signals truthfully and the principal can fully learn the unknown state of the world as the number of agents grows large. However, if the principal and the agents do not share the same preferences, that is, their interests conflict, then the agents might misrepresent information in their messages, as shown by Wolinsky (2002), Morgan and Stocken (2008), Levit and Malenko (2011), Battaglini (2017), and Ekmekci and Lauermann (2022) among others. Several of these studies show that if the conflict of interest between the principal and the agents is below a certain threshold, then as the number of agents grows large, the agents report their signals almost truthfully and the principal can still fully learn the unknown state, and if the conflict is above the threshold, the agents' messages become completely uninformative for any number of agents. However, the results in all of these cases depend on the critical assumption that the preferences of the principal and the agents are aligned if they have complete information about the realized state.

In many situations, the preferences of the principal and the agents might not be fully aligned even if they have complete information about the realized state. Consider the example of non-binding shareholder voting studied by Levit and Malenko (2011), in which the shareholders receive dispersed information concerning the unknown payoff of a proposal to the firm and decide whether to vote in favor of it, while the manager observes the outcome of the vote and ultimately forms his own decision. Both the manager and the shareholders care about the payoff of the proposal and agree on the same decision if the realized payoff is at the extremes of either very high or very low. However, if the manager receives additional private benefits from the proposal, then when the realized payoff is moderate, the preferences of the principal and the agents are more likely to be misaligned: in this case, only the manager might prefer the proposal due to his additional payoff. Similarly, in public protests studied by Battaglini (2017), the citizens receive dispersed information concerning the effect of reform and decide whether to participate in a rally, while the policymaker decides whether to implement the reform after observing the citizens' activities. The preferences of the politician and the citizens are aligned if the reform is dramatically better or worse than the *status-quo*, but are misaligned for less significant changes, where the policymaker's private interests or ideologies may play a larger role. A similar situation also arises in the example of survey polls.

In this paper, I model the conflict of interest between the principal and the agents more generally and consider the case where the preferences of the principal and the agents are misaligned in some realized states. I show that in the framework of the existing literature, if the conflict between the principal and the agents is moderate, there is a discontinuity: when the number of agents is large enough, adding even a tiny probability of misaligned states leads to complete unraveling in which the agents ignore their signals and no information is transmitted. This result stands in contrast to the predicted outcome in the existing literature, in which the agents report almost truthfully. In addition, I demonstrate that no matter how small the conflict of interest between the principal and the agents is, the information contained in each agent's message must vanish as the number of agents grows large. Finally, and more surprisingly, no matter how many agents there are, the total amount of information that is transmitted is limited, and the principal always fails to fully learn the unknown state.

More specifically, I develop a model based on Levit and Malenko (2011) and Battaglini (2017) (henceforth, LMB). Both of these papers analyze a model with one principal and N agents. The principal must decide between policy A and policy B. Both the principal and the agents find A optimal in the high state and B optimal in the low state, that is, their preferences are fully aligned when the state is known. All the agents have the same preferences, while the principal is biased toward A in each state. Therefore, when the realized state is uncertain, they have different "thresholds of acceptance": the principal already prefers A at a relatively low probability of the high state, while the agents prefer it only at a higher probability. For the information structure, both the principal and the agents share the same prior belief about the unknown state. Each agent has a small piece of private information about the realized state in the form of a noisy signal. She can then choose whether to approve A. The principal, in turn, observes the total number of approvals and then chooses a policy that is most in line with his interests.

LMB apply this type of model to non-binding shareholder voting and public protests, in which there are usually a large number of agents (shareholders or citizens). LMB show that information transmission is all-or-nothing. If the conflict of interest between the principal and the agents is below a certain threshold, then as N grows large, the agents report their signals almost truthfully. That is, they approve A with a probability approaching 1 when they receive signals favoring A, and they reject A with a probability approaching 1 when they receive signals opposing A. Hence, the principal can fully learn the unknown state, and the information dispersed among the agents is effectively aggregated. However, if the conflict is above the threshold, then complete unraveling happens, in which the agents ignore their signals, and in this case, no information is transmitted from the agents to the principal.

In what follows, I consider a further possibility, which can be exemplified with a relatively simple scenario. Let us add a middle state to LMB's framework.[1] This middle state is a misaligned state in which the principal prefers A while the agents prefer B. For the information structure, each agent's signal is ordered by the "monotone likelihood ratio property" (MLRP), which states that, as the realization of the signal increases, it becomes increasingly likely that the state is higher.

I show that when the conflict of interest in LMB's framework is moderate and below the threshold provided by LMB, there is a discontinuity in the results: when N is large enough, adding the middle state with even a tiny probability leads to complete unraveling in which the agents ignore their signals and no information is transmitted. This result stands in contrast to LMB's prediction, in which the agents report almost truthfully.

I demonstrate that when the conflict of interest in LMB's framework is sufficiently small, information is still transmitted from the agents to the principal if the middle state is also sufficiently unlikely. However, as the number of agents grows large, the information contained in an agent's message vanishes, that is, the agents reject A with a probability approaching 1 even when they receive the signal that favors A the most. Furthermore, the expected number of total approvals for A in each state is always smaller than a finite number that is independent of N. Similarly, the principal chooses A when the total number of approvals exceeds a cut-off number, and this cut-off is also always smaller than a finite number that is independent of N. Hence, the principal must follow either the unanimity rule under which he chooses A if at least one agent approves A or rules that are similar to the unanimity rule. Finally, I show that no matter how large N is, the total

1. The key difference between this paper and LMB is whether the preferences of the principal and agents are always fully aligned if they know the state, instead of whether there are two or three states.

amount of information that is transmitted is limited and the principal always fails to fully learn the unknown state. With a strictly positive probability, the principal chooses the wrong policy in both the high state and the low state, even though the preferences of the principal and the agents are fully aligned in both states.

Another important finding by Battaglini (2017) is that communication among the agents facilitates information transmission and aggregation, benefiting both the principal and the agents. Battaglini thus highlights the value of social media for the effectiveness of petitions and public protests since social media allow citizens to share information. In contrast, by further considering the case in which the agents fully communicate with each other, I show that communication among the agents might impede information transmission and hurt both the principal and the agents. In this case, as N approaches infinity, the agents learn the state, and information is effectively aggregated. However, I find that in some situations, the principal ignores messages from the agents if they fully communicate with each other, while if they cannot communicate, information transmission is restored. A key intuition is that we can interpret the failure of information aggregation as intentional vagueness that mitigates the conflict of interest between the sender (agents) and the receiver (principal), as discussed in the cheap-talk literature initiated by Crawford and Sobel (1982).

In this paper's basic model, the agents can either approve A or reject it, that is, they can only send binary messages. However, I also extend the model to the case where the set of available messages for the agents is not restricted to being binary, which allows the framework of this paper to capture some natural features of applications, for example, the possibilities of abstaining in non-binding shareholder voting, staying neutral in public protests, and sending medium scores in survey polls. In this case, the principal's decision rule becomes multi-dimensional rather than a cut-off in the total number of approvals, which complicates the analysis. I provide a novel and tractable way to analyze this case by taking inspiration from Chernoff's fundamental connection between simple statistical hypothesis tests and large deviation theory.[2] I show that all of the results presented above are robust in a natural class of equilibria in which the agents follow monotonic strategies.

It is also interesting to see how much information the principal can elicit if he can *ex-ante* commit to a decision rule. In LMB's framework, when N is large, the principal can approach his first-best outcome by committing to a voting mechanism with any qualified majority rule in which he chooses A if the ratio of approvals exceeds a certain cut-off. However, I show that in the present paper's framework, the principal cannot rely on any qualified majority rule since according to the Condorcet jury theorem, they all lead to the first-best outcome for the agents. However, the principal can approach his first-best outcome by randomizing

2. See Chernoff (1952)

between two qualified majority rules, that is, between two cut-offs in the ratio of approvals.

The rest of this paper proceeds as follows: Section 1.2 describes the model and characterizes the equilibrium. Section 1.3 presents the main result that learning is always incomplete no matter how many agents there are. Section 1.4 discusses information transmission from the agents to the principal. Section 1.5 analyzes the case where the set of available messages for the agents is not restricted to being binary. Section 1.6 studies the situation in which the principal can *ex-ante* commit to a decision rule. Section 1.7 surveys the related literature, and Section 1.8 concludes the paper. Most of the proofs are sketched in the main text, with the details relegated to the appendix.

1.2 Model

1.2.1 Basic Setting

There is one principal (he) and N agents (she). The principal has to decide between two policies, A and B. When he chooses B, the payoffs for all players are normalized to 0. When he chooses A, the payoffs for all players depend on an unknown state of the world $\theta \in \Theta$, with $\Theta = \{\theta_1, \ldots, \theta_n\} \subset \mathbb{R}$ and $\theta_1 < \ldots < \theta_n$.[3] In state θ, the principal receives the payoff $V_{pc}(\theta)$ by choosing A, while the agents all have the same preference and receive the payoff $V_{ag}(\theta)$.

Both the principal and the agents receive higher payoffs from A when the state is higher, that is, both $V_{pc}(\theta)$ and $V_{ag}(\theta)$ strictly increase with θ. There are thresholds $\hat{\theta}_{pc}, \hat{\theta}_{ag} \in \Theta$ such that for each $j \in \{pc, ag\}$:

$$\begin{aligned} V_j(\theta) &> 0 \quad \text{if} \quad \theta \geq \hat{\theta}_j, \\ V_j(\theta) &< 0 \quad \text{if} \quad \theta < \hat{\theta}_j. \end{aligned}$$

The principal prefers A more than the agents in every state, that is,[4]

$$V_{pc}(\theta) \geq V_{ag}(\theta), \ \forall \theta \in \Theta, \tag{1.1}$$

$$\hat{\theta}_{pc} < \hat{\theta}_{ag}. \tag{1.2}$$

3. In this paper, we mostly focus on the case where $n = 3$.

4. The condition (1.2) that the principal and the agents have different cut-offs for states is critical for results, while the condition (1.1) is for the better exposition. If (1.1) is violated, the main results of this paper (Theorems 1.1 and 1.2) still hold and the other results except for Lemma 1.1 can also be easily extended.

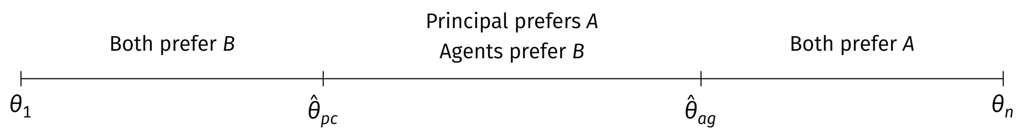

Figure 1.1. Preferences

Notes: The Preferences of the principal and the agents when the realized state is known. Their preferences are not aligned when $\theta \in [\hat{\theta}_{pc}, \hat{\theta}_{ag})$.

For the information structure, the principal and the agents share a common prior belief $q^0 = (q_1^0, ..., q_n^0) \in \Delta^n$ about the unknown state, with $q_j^0 > 0$ for each $j \in \{1, ..., n\}$. Conditional on the state $\theta \in \Theta$, each agent $i \in \{1, ..., N\}$ receives a private, independent signal $s^i \in \{\ell, h\}$, that is, a low or a high signal, with

$$\rho_j = \mathbb{P}[s^i = h|\theta_j], \ \forall \theta_j \in \Theta,$$

$$0 < \rho_1 < ... < \rho_n < 1. \tag{1.3}$$

Hence, the agents are more likely to receive signal h when the state is higher.

After observing the private signal, each agent chooses whether to approve A. The principal observes the total number of approvals $T \in \{0, ..., N\}$ and chooses the policy that maximizes his expected payoff.

1.2.2 Three-State Scenario

For simplicity, this paper focuses on the case where $\Theta = \{\theta_1, \theta_2, \theta_3\}$.[5] Both the principal and the agents prefer A to B in state θ_3 and prefer B to A in state θ_1:

$$V_{pc}(\theta_3) > 0 \text{ and } V_{ag}(\theta_3) > 0,$$

$$V_{pc}(\theta_1) < 0 \text{ and } V_{ag}(\theta_1) < 0,$$

while only the principal prefers A in state θ_2:

$$V_{pc}(\theta_2) > 0 \text{ and } V_{ag}(\theta_2) < 0, \tag{1.4}$$

that is, we have $\hat{\theta}_{pc} = \theta_2$ and $\hat{\theta}_{ag} = \theta_3$. The preferences of the principal and the agents are not aligned in state θ_2, which is a ***misaligned state***. The preferences of the principal and the agents are illustrated by the simplex of belief $q = (q_1, q_2, q_3) \in \Delta^3$ in Figure 1.2.

5. This setting allows us to incorporate and compare the results with the existing literature. The main results of this paper (Theorems 1.1 and 1.2) hold for the general setting in Section 1.2.1. The other results can also be easily extended.

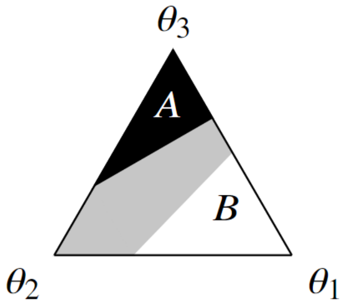

Figure 1.2. Preferences when the realized state is uncertain

Notes: The corner θ_i for $i \in \{1,2,3\}$ corresponds to the belief q with $q_i = 1$. The segment $\theta_i\theta_j$ corresponds to the set of beliefs $\{q|q_i + q_j = 1\}$. Both the principal and the agents prefer A when they all hold a belief q in the black area, while both prefer B in the white area. In the shaded area, only the principal prefers A.

If we assume that $q_2^0 = 0$ and ignore the misaligned state θ_2, then the preferences of the principal and the agents are aligned when the realized state is known and misaligned when the realized state is uncertain. The conflict of interest is generated by ***different payoff intensities in state $\boldsymbol{\theta_1}$ and $\boldsymbol{\theta_3}$*** between the principal and the agents. From (1.1):

$$V_{pc}(\theta_3) \geq V_{ag}(\theta_3) > 0 > V_{pc}(\theta_1) \geq V_{ag}(\theta_1).$$

Therefore,[6]

$$-\frac{V_{pc}(\theta_1)}{V_{pc}(\theta_3)} \leq -\frac{V_{ag}(\theta_1)}{V_{ag}(\theta_3)}. \tag{1.5}$$

As illustrated in Figure 1.3,[7] the principal and the agents have different thresholds of acceptance: for each belief $q = (q_1, q_3) \in \Delta^2$, the principal prefers A to B if $\frac{q_3}{q_1} > -\frac{V_{pc}(\theta_1)}{V_{pc}(\theta_3)}$, while the agents prefer A to B if $\frac{q_3}{q_1} > -\frac{V_{ag}(\theta_1)}{V_{ag}(\theta_3)}$.

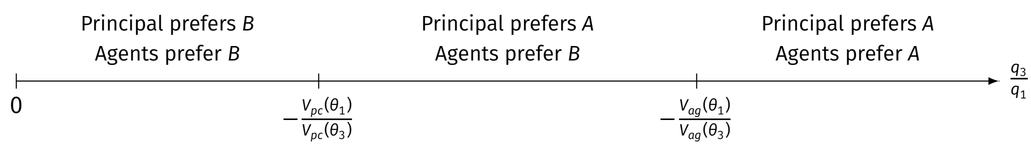

Figure 1.3. Different thresholds of acceptance when $q_2 = 0$.

6. All results hold if (1.5) is valid but (1.1) is violated. When (1.2) is valid, the condition (1.5) is equivalent to the argument that under each belief, if the agents prefer A, the principal does so. If (1.5) is violated, the main results of this paper (Theorems 1.1 and 1.2) still hold and the other results except for Lemma 1.1 can also be easily extended.

7. Note that when we assume $q_2^0 = 0$ and ignore the misaligned state θ_2, we suppress Figure 1.2 to its $\theta_3\theta_1$ segment, which we convert to Figure 1.3.

1.2.3 Strategy and Equilibrium

We examine the symmetric *Bayesian Nash equilibrium*, in which all the agents use the same strategy $\boldsymbol{x} = (x_\ell, x_h) \in [0, 1]^2$. Each agent $i \in \{1, ..., N\}$ approves A with probabilities x_ℓ and x_h respectively, when $s^i = \ell$ and $s^i = h$.

We consider equilibria in which the agents who receive signal h are more likely to approve A than the agents who receive signal ℓ, that is, $x_\ell \le x_h$.[8] Note that there always exists a babbling equilibrium in which $x_\ell = x_h$, that is, the agents ignore their signals. In this equilibrium, the principal finds the total number of approvals uninformative and makes a decision based only on his prior belief. There is complete unraveling, and no information is transmitted from the agents to the principal.

We now consider the case where $x_\ell < x_h$. In this case, the principal forms his posterior belief based on his prior belief and the total number of approvals T. The posterior likelihood ratios[9]

$$\frac{\mathbb{P}[T; N|\theta_3]}{\mathbb{P}[T; N|\theta_1]} \text{ and } \frac{\mathbb{P}[T; N|\theta_2]}{\mathbb{P}[T; N|\theta_1]}$$

strictly increase with T since the agents are more likely to receive higher signals and hence are more likely to approve A when the realized state is higher, that is,

$$\rho_1 x_h + (1-\rho_1)x_\ell \ < \ \rho_2 x_h + (1-\rho_2)x_\ell \ < \ \rho_3 x_h + (1-\rho_3)x_\ell,$$

by (1.3) and $x_\ell < x_h$. Thus, the principal's posterior belief that the realized state is θ_1 strictly decreases with T.

Hence, a pure strategy for the principal is a cut-off $\hat{T} \in \{0, ..., N+1\}$ such that he chooses A if and only if $T \ge \hat{T}$. A mixed strategy for the principal allows him to randomize when he observes $\hat{T}$ approvals. For simplicity, we assume in the main text that the principal always chooses B when he is indifferent and hence that he always uses pure strategies. All results remain valid when the principal can use a mixed strategy, as shown in the appendix.[10]

We focus on the ***informative equilibrium*** in which the agents use an ***informative strategy*** $\boldsymbol{x}$ with $x_\ell < x_h$, that is, the agents make decisions according to

8. For equilibria with $x_\ell \ge x_h$, we relabel approving A as rejecting A, and the following analyses still hold.

9. $\mathbb{P}[T; N|\theta_i] = \binom{N}{T}[\underbrace{\rho_i x_h + (1-\rho_i)x_\ell}_{\text{Prob of approving } A}]^T[\underbrace{1-\rho_i x_h - (1-\rho_i)x_\ell}_{\text{Prob of rejecting } A}]^{N-T}$, $\forall i \in \{1, 2, 3\}$.

10. In particular, we do not rely on the principal's mixed strategies to ensure the existence of informative equilibria. In the appendix, we characterize the equilibria where the principal can use a mixed strategy and show that if there exists an equilibrium in which the principal randomizes at $\hat{T} \in \{0, ..., N-1\}$, then there must exist an equilibrium in which the principal chooses A if and only if $T > \hat{T}$. If there exists an equilibrium in which the principal randomizes at $\hat{T} = N$, then there exists an equilibrium in which the principal chooses A if and only if $\hat{T} = N$.

their private information, and the principal uses a ***responsive strategy*** $\hat{T}$ with $\hat{T} \in \{1, ..., N\}$, that is, the principal makes decisions according to the number of approvals.

1.2.4 Characterization of Informative Equilibria

<u>Best Response of the Agents</u>: Consider a strategy profile in which the agents choose an informative strategy $\boldsymbol{x}$ and the principal chooses a responsive strategy $\hat{T}$. An agent is *pivotal* if the principal receives $\hat{T}-1$ approvals from the other $N-1$ agents. When deciding whether to approve A, it is optimal for an agent to condition on the pivotal event since this agent's decision cannot affect the outcome in any other event. The likelihood of being pivotal in state $\theta_i \in \{\theta_1, \theta_2, \theta_3\}$ is

$$\mathbb{P}[piv|\theta_i; \boldsymbol{x}, \hat{T}] = \binom{N-1}{\hat{T}-1} [\underbrace{\rho_i x_h + (1-\rho_i) x_\ell}_{\text{Prob of approving } A}]^{\hat{T}-1} [\underbrace{1 - \rho_i x_h - (1-\rho_i) x_\ell}_{\text{Prob of rejecting } A}]^{N-\hat{T}}.$$

When this agents receives $s \in \{\ell, h\}$, she approves A only if

$$\sum_{i=1}^{3} \underbrace{q_i^0}_{\text{prior}} \cdot \underbrace{\mathbb{P}[s|\theta_i]}_{signal} \cdot \underbrace{\mathbb{P}[piv|\theta_i; \boldsymbol{x}, \hat{T}]}_{\text{being pivotal}} \cdot V_{ag}(\theta_i) \geq 0.$$

Rewriting this as a payoff-weighted likelihood ratio, we have

$$\frac{q_3^0 \cdot \mathbb{P}[s|\theta_3] \cdot \mathbb{P}[piv|\theta_3; \boldsymbol{x}, \hat{T}] \cdot V_{ag}(\theta_3)}{-\sum_{i=1}^{2} q_i^0 \cdot \mathbb{P}[s|\theta_i] \cdot \mathbb{P}[piv|\theta_i; \boldsymbol{x}, \hat{T}] \cdot V_{ag}(\theta_i)} \geq 1.$$

Letting $L_{ag}(s; \boldsymbol{x}, \hat{T})$ denote the left side. This agent chooses $\boldsymbol{x}$ as the best response if for each $s \in \{\ell, h\}$,

$$\begin{cases} x_s = 1 & \text{when } L_{ag}(s; \boldsymbol{x}, \hat{T}) > 1, \\ x_s \in [0,1] & \text{when } L_{ag}(s; \boldsymbol{x}, \hat{T}) = 1, \\ x_s = 0 & \text{when } L_{ag}(s; \boldsymbol{x}, \hat{T}) < 1. \end{cases} \tag{1.6}$$

By (1.3), we have

$$L_{ag}(h; \boldsymbol{x}, \hat{T}) > L_{ag}(\ell; \boldsymbol{x}, \hat{T}). \tag{1.7}$$

By (1.6) and (1.7), if $\boldsymbol{x}$ with $x_\ell < x_h$ is the best response to itself and $\hat{T} \in \{1, ..., N\}$, then it must satisfy the following:

$$\begin{cases} x_h = 1 & \text{if } x_\ell > 0, \\ x_\ell = 0 & \text{if } x_h < 1. \end{cases}$$

Best Response of the Principal: Consider the case where the agents choose an informative strategy $\boldsymbol{x}$. When the principal observes T approvals from N agents, he chooses A only if

$$\sum_{i=1}^{3} \underbrace{q_i^0}_{\text{prior}} \cdot \underbrace{\mathbb{P}[T;N|\theta_i]}_{T \text{ approvals}} \cdot V_{pc}(\theta_i) \;>\; 0.$$

Rewriting this as a payoff-weighted likelihood ratio, we have

$$\frac{\sum_{i=2}^{3} q_i^0 \cdot \mathbb{P}[T;N|\theta_i] \cdot V_{pc}(\theta_i)}{-q_1^0 \cdot \mathbb{P}[T;N|\theta_1] \cdot V_{ag}(\theta_1)} \;>\; 1.$$

Letting $L_{pc}(T;\boldsymbol{x})$ denote the left side.[11] Note that $L_{pc}(T;\boldsymbol{x})$ strictly increases with T since

$$\rho_1 x_h + (1-\rho_1)x_\ell \;<\; \rho_2 x_h + (1-\rho_2)x_\ell \;<\; \rho_3 x_h + (1-\rho_3)x_\ell$$

by (1.3) and $x_\ell < x_h$, that is, the agents are more likely to approve A when the realized state is higher. The optimal cut-off for the principal is

$$\hat{T} \;=\; \min\{T \mid L_{pc}(T;\boldsymbol{x}) \;>\; 1 \text{ and } T \;\in\; \{0,\ldots,N+1\}\}. \tag{1.8}$$

Informative Equilibrium: An ***informative equilibrium*** is characterized by a pair $\{(x_\ell, x_h), \hat{T}\}$ that satisfies (1.6) and (1.8), with $x_\ell < x_h$ and $\hat{T} \in \{1,\ldots,N\}$.

We can show that in every informative equilibrium, the agents always reject A when they receive signal ℓ:

Lemma 1.1. *The agents choose $x_\ell = 0$ in every informative equilibrium.*

For a sketch of the proof, consider an informative equilibrium with $x_\ell > 0$. From (1.6) and (1.7), we have $x_h = 1$ and hence $x_\ell \in (0,1)$. Therefore, conditional on being pivotal, the agents always approve A when they receive signal h and are indifferent between A and B when they receive signal ℓ. Note that the principal prefers A more than the agents do. If an agent is indifferent conditional on being pivotal and receiving signal ℓ, then the principal prefers A when this agent is pivotal and receives signal ℓ. Since this agent only rejects A when she receives signal ℓ, the principal prefers A when this agent is pivotal and rejects A, that is, when the principal observes $\hat{T}-1$ approvals. However, it leads to a contradiction to the optimality of $\hat{T}$ as shown in (1.8).

11. We suppress $\boldsymbol{x}$ in $\mathbb{P}[T;N|\theta_i]$ for each $i \in \{1,2,3\}$.

1.3 Information Aggregation

In many situations, including non-binding shareholder voting, public protests, and survey polls, among others, there are usually a large number of agents (shareholders, citizens, and interviewees). In this section, we study whether information dispersed among the agents is effectively aggregated and whether the principal fully learns the state as the number of agents grows large:

Definition 1.1. A sequence of equilibria $\{\Gamma_N\}_{N=1}^{\infty}$ **aggregates information** if

$$\lim_{N\to\infty} \mathbb{P}[A|\theta_1;\Gamma_N] + \mathbb{P}[B|\theta_3;\Gamma_N] \; = \; 0.$$

We consider information aggregation with minimal requirements by focusing on stated θ_1 and θ_3 in which the preferences of the principal and the agents are aligned. Note that the failure of information aggregation implies that the principal fails to fully learn the state no matter how many agents there are.

The present paper's framework shares certain qualitative features with elections in which voters decide whether to approve a policy and the total number of approvals matters. As shown by the Condorcet jury theorem (see Ladha (1992)) and its modern versions (Feddersen and Pesendorfer (1997), Feddersen and Pesendorfer (1998), Myerson (1998), Duggan and Martinelli (2001)), elections effectively aggregate dispersed information among the agents (voters) under any qualified majority rule that depends on the ratio of votes. However, full information aggregation fails under the unanimity rule or rules that are close to it.

The fundamental difference between this paper's framework and elections is that the principal can now choose the policy based on his own decision and is not constrained by predetermined rules. The existing literature extends the idea behind the Condorcet jury theorem and shows that information is still effectively aggregated if the conflict of interest between the principal and the agents is small. However, we show that full information aggregation always fails after adding the misaligned state θ_2.

1.3.1 Results from the Existing Literature

In this section, we assume that $q_2^0 = 0$ and ignore the misaligned state θ_2. As discussed in Section 1.2.2, the preferences of the principal and the agents are fully aligned if the realized state is known and misaligned when the state is uncertain. The principal and the agents have different thresholds of acceptance: for each belief $q = (q_1, q_3) \in \Delta^2$, the principal prefers A if $\frac{q_3}{q_1} > -\frac{V_{pc}(\theta_1)}{V_{pc}(\theta_3)}$, while the agents prefer A if $\frac{q_3}{q_1} > -\frac{V_{ag}(\theta_1)}{V_{ag}(\theta_3)}$. Hence, the ratio of $-\frac{V_{ag}(\theta_1)}{V_{ag}(\theta_3)}$ to $-\frac{V_{pc}(\theta_1)}{V_{pc}(\theta_3)}$ is a natural measure for the conflict of interest between the principal and the agents due to the different payoff intensities in state θ_1 and state θ_3. The existing literature has considered this case and shown that if the ratio $\frac{V_{ag}(\theta_1)}{V_{ag}(\theta_3)}\frac{V_{pc}(\theta_3)}{V_{pc}(\theta_1)}$ is below a certain threshold, information is effectively aggregated.[12]

Proposition 1.1. *Assume that $q_2^0 = 0$.*

1. If

$$\frac{V_{ag}(\theta_1)}{V_{ag}(\theta_3)} \cdot \frac{V_{pc}(\theta_3)}{V_{pc}(\theta_1)} < \frac{\rho_3}{\rho_1} \cdot \frac{1-\rho_1}{1-\rho_3}, \tag{1.9}$$

then there exists a sequence of equilibria that aggregates information.

2. If

$$\frac{V_{ag}(\theta_1)}{V_{ag}(\theta_3)} \cdot \frac{V_{pc}(\theta_3)}{V_{pc}(\theta_1)} > \frac{\rho_3}{\rho_1} \cdot \frac{1-\rho_1}{1-\rho_3}, \tag{1.10}$$

then only the babbling equilibrium exists for each N.

Figure 1.4 illustrates the intuition behind Proposition 1.1. In an informative equilibrium, the agents who receive signal h must (weakly) prefer A conditional on being pivotal. That is, each agent's posterior belief must be higher than $-\frac{V_{ag}(\theta_1)}{V_{ag}(\theta_3)}$ conditional on signal h and $\hat{T}-1$ approvals from the other $N-1$ agents. However, the principal optimally chooses the cut-off $\hat{T}$. He must prefer B when he observes $\hat{T}-1$ approvals from N agents. That is, his posterior belief must be less than $-\frac{V_{pc}(\theta_1)}{V_{pc}(\theta_3)}$ when there are already $\hat{T}-1$ approvals from $N-1$ agents and the pivotal agent rejects A. Therefore, the difference in the thresholds of the posterior likelihood ratio between the agents and the principal depends at most on one signal h and one rejection in every informative equilibrium.

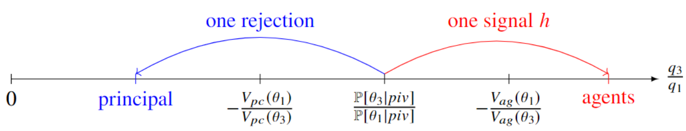

Figure 1.4. Inference from being pivotal and thresholds of acceptance

Notes: The red line corresponds to the argument that the agents signal h must prefer A conditional on being pivotal. The blue line corresponds to the argument that the principal must prefer B when he observes $\hat{T}-1$ approvals, that is, when the pivotal agent rejects A.

Note that if the agents report their signals truthfully, that is, if the agents approve A when they receive signal h and reject A when they receive signal ℓ, the decrease in the posterior likelihood ratio due to one rejection is maximized. Hence, we can replace "one rejection" in Figure 1.4 with "one signal ℓ" and argue

12. Levit and Malenko (2011) consider the setting with a symmetric information structure. Battaglini (2017) considers the setting in which the number of voters follows a Poisson distribution. Ekmekci and Lauermann (2022) consider the setting with deterministic population size in their online appendix. We sketch their proof.

that a necessary condition for the existence of informative equilibria is that the difference in the thresholds of the posterior likelihood ratio depends on at most one signal h and one signal ℓ, and thus we derive (1.9).

The inequality (1.9) is indeed a necessary condition for the existence of the informative equilibrium in which the agents report truthfully, that is, it is a necessary condition for the existence of $\hat{T}$ such that

$$\frac{\mathbb{P}[\hat{T};N|\theta_3]}{\mathbb{P}[\hat{T};N|\theta_1]} \geq -\frac{V_{ag}(\theta_1)}{V_{ag}(\theta_3)},$$
$$\frac{\mathbb{P}[\hat{T}-1;N|\theta_3]}{\mathbb{P}[\hat{T}-1;N|\theta_1]} \leq -\frac{V_{pc}(\theta_1)}{V_{pc}(\theta_3)},$$

when $x_h = 1$ and $x_\ell = 0$. However, it is not a sufficient condition due to the requirement that $\hat{T}$ must be an integer. We show that as N grows large, the effect of this integer requirement vanishes, and there exists an informative equilibrium in which the agents report almost truthfully, with $x_h \approx 1$ and $x_\ell = 0$:

Lemma 1.2. *Assume that $q_2^0 = 0$. If $\frac{V_{ag}(\theta_1)}{V_{ag}(\theta_3)}\frac{V_{pc}(\theta_3)}{V_{pc}(\theta_1)} < \frac{\rho_3}{\rho_1}\frac{1-\rho_1}{1-\rho_3}$, then for each ϵ, there exists an N_ϵ such that for each $N > N_\epsilon$, there exists an informative equilibrium in which the agents choose $x_h = 1-\epsilon$ and $x_\ell = 0$ (almost truthtelling).*

Therefore, when the conflict generated by the different payoff intensities is below the threshold, the principal can fully learn the unknown state by the law of large numbers as $N \to \infty$, and information is effectively aggregated.

1.3.2 Failure of Information Aggregation

We consider the setting with the misaligned state θ_2 by assuming that $q_2^0 > 0$ in the rest of this paper. In this setting, the preferences of the principal and the agents might be misaligned even if they know the realized state.

As reviewed in Section 1.3.1, the existing literature provides the condition under which informative equilibria exist and shows that if this condition is satisfied, there exists a sequence of informative equilibria that aggregates information with the agents reporting their signals almost truthfully as $N \to \infty$. We now show that when $q_2^0 > 0$, full information aggregation always fails even if informative equilibria exist. The principal, therefore, fails to fully learn the realized state even if he receives a large number of informative messages.

Theorem 1.1. *No sequence of equilibria aggregates information. That is, there exists a constant $c > 0$[13] such that for each N and each equilibrium Γ with N agents,*

$$\mathbb{P}[A|\theta_1;\Gamma] + \mathbb{P}[B|\theta_3;\Gamma] > c.$$

13. The value of c depends on other parameters, as do N_ϵ, T_0, $\hat{T}^*$, N^*, M_1, and δ, introduced later in this section.

We illustrate the intuition behind Theorem 1.1 through two steps.

Step 1. Vanishing Information

We first argue that when $q_2^0 > 0$, the information contained in an agent's message must vanish as $N \to \infty$ in every sequence of informative equilibria, which differs sharply from Lemma 1.2. We also show how quickly the information vanishes, that is, the rate of convergence for $x_h \to 0$.

Proposition 1.2. *For each $\epsilon > 0$, there exists N'_ϵ such that when $N > N'_\epsilon$, the agents choose $x_h < \epsilon$ in every informative equilibrium. Furthermore, there exists $T_0 > 0$ such that for each N and each informative equilibrium with N agents,*

$$N \cdot x_h \; < \; T_0.$$

To understand the intuition, fix an arbitrary $x \in (0,1])$ and suppose that the agents behave according to $x_h = x$ and $x_\ell = 0$ for each $N \in \mathbb{N}^+$. The expected number of approvals in each state θ_i is $N \cdot \rho_i x$ for $i \in \{1,2,3\}$. Figure 1.5 illustrates the distributions of the total number of approvals T when N is large.[14] We can see that when N is large, (i) the principal chooses $\hat{T}$ such that $N\rho_1 x < \hat{T} < N\rho_2 x$ since he prefers A when the realized state is θ_2 or θ_3, and (ii) as shown in Figure 1.5,

$$\frac{\mathbb{P}[\hat{T};N|\theta_3]}{\mathbb{P}[\hat{T};N|\theta_2]} \approx 0,$$

and hence

$$\frac{\mathbb{P}[piv|\theta_3]}{\mathbb{P}[piv|\theta_2]} = \frac{\mathbb{P}[\hat{T}-1;N-1|\theta_3]}{\mathbb{P}[\hat{T}-1;N-1|\theta_2]} \approx \frac{\mathbb{P}[\hat{T};N|\theta_3]}{\mathbb{P}[\hat{T};N|\theta_2]} \approx 0.$$

Each agent believes that the realized state is very unlikely to be θ_3 conditional on being pivotal. She rejects A even when she receives signal h. Hence, she does not choose $x_h = x$ as a best response.

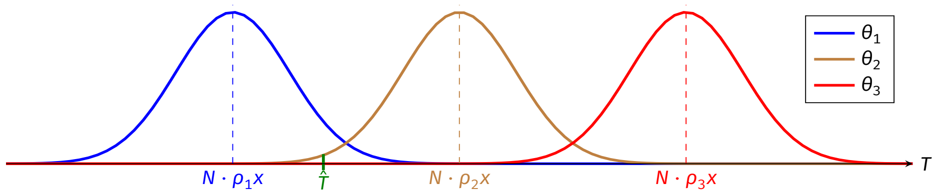

Figure 1.5. The distributions of the total number of approvals

Notes: The distribution of the total number of approvals for A in each state when the agents choose $x_h = x \in (0,1)$ and $x_\ell = 0$. The principal optimally chooses $\hat{T}$.

14. Note that we can approximate these distributions by normal distributions.

Each agent makes her decision conditional on her signal and being pivotal. However, the number of approvals that makes an agent pivotal is endogenous, since the principal must be nearly indifferent between A and B when he observes this number. If the agents' messages are informative and the principal receives a large number of messages, then each agent believes that the realized state must be either θ_1 or θ_2 given that the principal is uncertain whether the realized state is θ_1 and indifferent between A and B, as shown in Figure 1.5. Hence, each agent prefers B regardless of her signal, conditional on being pivotal.

To prevent the inference conditional on being pivotal from overwhelming each agent's private information, the distributions of the total number of approvals in different states must be close to each other, as shown in Figure 1.6.[15,16] Hence, the information contained in an agent's message must vanish as $N \to \infty$. We then show that the information in an agent's message must vanish at a high speed to make the differences in the mean $N \cdot \rho_i x_h$ of different states finite. Thus, the rate of convergence for $x_h \to 0$ must be comparable to $\frac{1}{N}$.

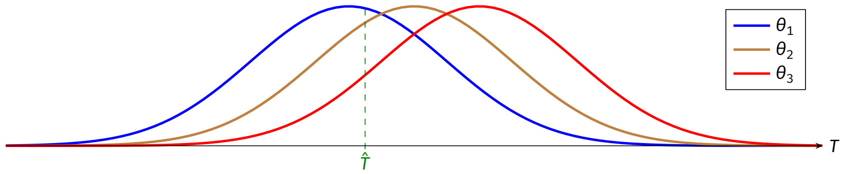

Figure 1.6. The distributions of the total number of approvals

Notes: The distributions of the total number of approvals in different states must be close to each other.

For an alternative intuition behind Proposition 1.2, once again fix an arbitrary $x \in (0, 1]$ and suppose that the agents behave according to $x_h = x$ and $x_\ell = 0$. In this case, the principal and the agents have different preferences, that is, they have different cut-offs for the total number of approvals above which A should be implemented. If the difference in cut-offs is large, the strategy profile with $x_h = x$ and $x_\ell = 0$ cannot be a part of an informative equilibrium. However, the difference in cut-offs increases with x, which measures the information in an agent's message, and the number of agents N. Therefore, the information contained in an agent's message must vanish as $N \to \infty$ in every sequence of informative equilibria. Note that in LMB's setting analyzed in Section 1.3.1, the difference in cut-offs is constant with respect to N and decreases with x if we ignore the integer requirement for $\hat{T}$.

15. When $x_\ell = 0$ and Nx_h is finite, we can approximate the distributions of the total number of approvals by Poisson distributions.

16. One might directly see the failure of information aggregation from Figure 1.6. The distribution in state θ_1 must be close to the distribution in state θ_3. Hence, the principal cannot find a $\hat{T}$ to separate them.

Step 2. Unanimity Rule

By Proposition 1.2, the expected number of approvals in each state is always smaller than a finite number that is independent of N. We also show that the principal's cut-off $\hat{T}$ is always smaller than a finite number that is independent of N. Hence, the principal must follow either the unanimity rule ($\hat{T} = 1$) such that he chooses B only if all the agents reject A or rules that are similar to the unanimity rule.

Proposition 1.3. *There exists $T_0 > 0$ such that for each N and each informative equilibrium with N agents,*

$$\hat{T} < T_0.$$

Note that for each N and each informative equilibrium with N agents, the posterior beliefs about state θ_3 and state θ_1 must have the same magnitude conditional on being pivotal, that is, there exists an $M_1 > 0$ such that

$$\frac{1}{M_1} < \frac{\mathbb{P}[\theta_3|piv]}{\mathbb{P}[\theta_1|piv]} < M_1. \tag{1.11}$$

If $\frac{\mathbb{P}[\theta_3|piv]}{\mathbb{P}[\theta_1|piv]} \to 0$, then the agents believe that the realized state is either θ_1 or θ_2 conditional on being pivotal, and they reject A when they receive $s = h$. If $\frac{\mathbb{P}[\theta_3|piv]}{\mathbb{P}[\theta_1|piv]} \to \infty$, then the principal believes that the realized state is either θ_2 or θ_3 when he observes $\hat{T} - 1$ approvals, and then chooses A.

In what follows, we show that there is no sequence of equilibria that satisfies

$$\lim_{N\to\infty} \sum_{T=0}^{\hat{T}-1} \mathbb{P}[T;N|\theta_1] = 1 \text{ and } \lim_{N\to\infty} \sum_{T=0}^{\hat{T}-1} \mathbb{P}[T;N|\theta_3] = 0,$$

since (i) $\mathbb{P}[\hat{T}-1;N|\theta_1]$ and $\mathbb{P}[\hat{T}-1;N|\theta_3]$ have the same magnitude as shown in (1.11), and (ii) $\hat{T}$ is always smaller than a finite number T_0 as shown in Proposition 1.3. Note that the left term is the probability that the principal chooses B in state θ_1, while the right term is the probability that the principal chooses B in state θ_3. Therefore, no sequence of equilibria aggregates information. Further, we can show that the principal chooses the wrong policy with a strictly positive probability in each state:

Corollary 1.1. *There exists $\bar{\delta} > 0$ such that for each N and each informative equilibrium with N agents,*

$$\mathbb{P}[A|\theta] > \bar{\delta}, \ \forall \theta \in \{\theta_1, \theta_2, \theta_3\},$$
$$\mathbb{P}[B|\theta] > \bar{\delta}, \ \forall \theta \in \{\theta_1, \theta_2, \theta_3\}.$$

1.4 Information Transmission

In this section, we first discuss conditions under which informative equilibria exist. If there exist multiple informative equilibria, we can rank them in both the Blackwell order and the Pareto order. Hence, we can identify the most informative equilibrium that also maximizes the payoffs of the principal and the agents. We then discuss the amount of information transmission by focusing on the most informative equilibrium and show that the amount of information transmission decreases with the conflict of interest between the principal and the agents. Finally, we argue that it might be better to disperse information among the agents instead of letting one agent receive all the information and further argue that communication among the agents might impede information transmission and hurt both the principal and the agents.

1.4.1 Existence of Informative Equilibria

We say ***information transmission persists*** if there exists N_1 such that for each $N > N_1$, an informative equilibrium exists. We say ***information transmission fails*** if there exists N_2 such that for each $N > N_2$, only the babbling equilibrium exists.

Proposition 1.1 indicates that when $q_2^0 = 0$, information transmission fails if

$$\frac{V_{ag}(\theta_1)}{V_{ag}(\theta_3)} \cdot \frac{V_{pc}(\theta_3)}{V_{pc}(\theta_1)} > \frac{\rho_3}{\rho_1} \cdot \frac{1-\rho_1}{1-\rho_3}.$$

We now provide a new condition when $q_2^0 > 0$:

Proposition 1.4. *When $q_2^0 > 0$, information transmission fails if*

$$\frac{V_{ag}(\theta_1)}{V_{ag}(\theta_3)} \cdot \frac{V_{pc}(\theta_3)}{V_{pc}(\theta_1)} > \frac{\rho_3}{\rho_1}.$$

Therefore, when $\frac{V_{ag}(\theta_1)}{V_{ag}(\theta_3)} \frac{V_{pc}(\theta_3)}{V_{pc}(\theta_1)} \in (\frac{\rho_3}{\rho_1}, \frac{\rho_3}{\rho_1} \frac{1-\rho_1}{1-\rho_3})$, there exists a sequence of equilibria that aggregates information if $q_2^0 = 0$ from Proposition 1.1, while information transmission fails if $q_2^0 > 0$, that is, only the babbling equilibrium exists when N is large enough.

When we assume $q_2^0 = 0$ and ignore the misaligned state θ_2, we show that informative equilibria exist if the difference in the thresholds of the posterior likelihood ratio between the principal and the agents depends at most on one signal h and one rejection. We then let the agents report messages sincerely to maximize the information contained in one rejection and hence in one message.

However, the agents cannot send messages sincerely when $q_2^0 > 0$ and N is large, as discussed before. Otherwise, the agents infer that the state must be either θ_1 or θ_2 conditional on being pivotal, and ignore their signals. Instead, they choose $x_h \approx 0$ and their messages are nearly uninformative according to Proposition 1.2. Therefore, when N is large, in an informative equilibrium, the agents signal h are indifferent between A and B conditional on being pivotal, that is, the posterior likelihood ratio of state θ_3 to state θ_1 conditional on signal h and $\hat{T}-1$ approvals from $N-1$ agents must be higher than $-\frac{V_{ag}(\theta_1)}{V_{ag}(\theta_3)}$. However, since the principal prefers B when he observes $\hat{T}-1$ approvals from N agents, the posterior likelihood ratio of θ_3 to state θ_1 conditional on $\hat{T}-1$ approvals from N agents must be lower than $-\frac{V_{pc}(\theta_1)}{V_{pc}(\theta_3)}$. Then since almost no information is contained in an agent's message hence in one rejection, the posterior likelihood ratio conditional on $\hat{T}-1$ approvals from $N-1$ agents must also be lower than $-\frac{V_{pc}(\theta_1)}{V_{pc}(\theta_3)}$. Therefore, the difference in the thresholds of the posterior likelihood ratio $-\frac{V_{ag}(\theta_1)}{V_{ag}(\theta_3)}$ and $-\frac{V_{pc}(\theta_1)}{V_{pc}(\theta_3)}$ depends on at most one signal h.

Note that when $\frac{V_{ag}(\theta_1)}{V_{ag}(\theta_3)}\frac{V_{pc}(\theta_3)}{V_{pc}(\theta_1)} \in (\frac{\rho_3}{\rho_1}, \frac{\rho_3}{\rho_1}\frac{1-\rho_1}{1-\rho_3})$ and $q_2^0 > 0$, informative equilibria might exist when N is small. In this situation, the agents can choose an x_h away from 0, which increases the information contained in one rejection and hence makes up for a larger difference in the thresholds of the posterior likelihood ratio. Note that both the principal and the agents receive higher expected payoffs from any informative equilibria than from the babbling equilibrium. Hence, the amount of information transmission and the welfare for the principal and the agents[17] are not monotonic with respect to N. The expected payoffs of both the principal and the agents are maximized if the number of agents equals some finite number. In contrast, both the principal and the agents receive a lower expected payoff when we let the number of agents go to infinity. We discuss the effect of the number of agents on the welfare more generally in Section 1.8.

When $\frac{V_{ag}(\theta_1)}{V_{ag}(\theta_3)}\frac{V_{pc}(\theta_3)}{V_{pc}(\theta_1)} < \frac{\rho_3}{\rho_1}$, we argue that information transmission persists if the misaligned state θ_2 is unlikely, that is, the prior q_2^0 is small. However, we cannot freely vary q_2^0 due to the constraint that $q^0 \in \Delta^3$. We replace q^0 with $\lambda = \{\lambda_1, \lambda_2\}$ such that

$$\lambda_1 = \frac{q_1^0}{q_3^0} \quad \text{and} \quad \lambda_2 = \frac{q_2^0}{q_3^0}.$$

The ratio λ_2 measures the conflict of interest between the principal and the agents concerning the misaligned state θ_2. Both q_1^0 and q_3^0 are smaller while q_2^0 is larger when λ_2 is larger and λ_1 is constant.

Proposition 1.5. *If $\frac{V_{ag}(\theta_1)}{V_{ag}(\theta_3)}\frac{V_{pc}(\theta_3)}{V_{pc}(\theta_1)} < \frac{\rho_3}{\rho_1}$, then there exists $\hat{\lambda}_2 > 0$ such that*[18]

17. In Section 1.4.3, we show that if there exists at least one informative equilibrium, there exists an informative equilibrium that maximizes the amount of information transmission and the welfare of the principal and the agents among all informative equilibria.

18. The value of $\hat{\lambda}_2$ depends on the value of other parameters except λ_2 and N.

(1) if $\lambda_2 < \hat{\lambda}_2$, information transmission persists,

(2) if $\lambda_2 > \hat{\lambda}_2$, information transmission fails.

In the appendix, we provide sufficient and necessary conditions under which there exists an informative equilibrium with $\hat{T} = 1$, that is, the equilibrium in which the principal chooses the unanimity rule. We show that there exists $\hat{\lambda}_{2,1}$ such that when N is large, an informative equilibrium with $\hat{T} = 1$ exists if $\lambda_2 < \hat{\lambda}_{2,1}$ and only if $\lambda_2 \leq \hat{\lambda}_{2,1}$. We then extend this approach and further derive $\hat{\lambda}_{2,j}$ for each $j \in \mathbb{N}$ corresponding to the informative equilibrium with $\hat{T} = j$ and derive[19]

$$\hat{\lambda}_2 = \sup_{j \in \mathbb{N}} \hat{\lambda}_{2,j}.$$

We can further show that

$$\lim_{j \to \infty} \hat{\lambda}_{2,j} = 0.$$

Hence, for each $\lambda_2 > 0$, there exists T^* that is independent of N such that

$$\lambda_2 > \hat{\lambda}_{2,j}, \ \forall j > T^*,$$

which is also indicated by Proposition 1.3 that $\hat{T}$ in any informative equilibrium is always smaller than a number that is independent of N. The principal must follow the unanimity rule or rules that are close to it, which leads to the failure of information aggregation.

We now investigate how $\hat{\lambda}_2$ changes with other parameters.

Corollary 1.2. *The threshold $\hat{\lambda}_2$ increases with $V_{ag}(\theta_i)$ and decreases with $V_{pc}(\theta_i)$ for each $i \in \{1, 2, 3\}$.* [20]

From Corollary 1.2, the threshold $\hat{\lambda}_2$ decreases with $\frac{V_{ag}(\theta_1)}{V_{ag}(\theta_3)} \frac{V_{pc}(\theta_3)}{V_{pc}(\theta_1)}$ when we only vary one term. Note that $\hat{\lambda}_2$ measures the conflict of interest concerning the misaligned state θ_2 while $\frac{V_{ag}(\theta_1)}{V_{ag}(\theta_3)} \frac{V_{pc}(\theta_3)}{V_{pc}(\theta_1)}$ measures the conflict of interest concerning the payoff intensities in state θ_1 and state θ_3. Therefore, information transmission persists if both types of conflict are small, as shown by Figure 1.7.

19. When $\lambda_2 > \hat{\lambda}_2$, we only need to consider informative equilibria with $\hat{T} < T_0$ by Proposition 1.3 that guarantees that there exists N_2 above which no informative equilibrium exists.

20. When discussing the comparative statistics in Corollary 1.2, Corollary 1.3, and Corollary 1.4, we always change one parameter and keep the others fixed.

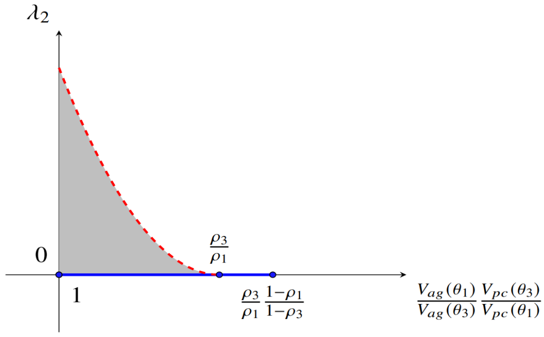

Figure 1.7. Information transmission and aggregation

Notes: The existing literature considers the case where $\lambda_2 = 0$, and shows that there exists a sequence of equilibria that transmits and aggregates information when $\frac{V_{ag}(\theta_1)}{V_{ag}(\theta_3)}\frac{V_{pc}(\theta_3)}{V_{pc}(\theta_1)} < \frac{\rho_3}{\rho_1}\frac{1-\rho_1}{1-\rho_3}$. When $\lambda_2 > 0$, we show that information transmission persists in the shaded area while information aggregation always fails.

Corollary 1.3. *The threshold $\hat{\lambda}_2$ decreases with ρ_1 while it is not monotonic with ρ_2 and ρ_3.*

Consider the case where ρ_2 decreases. On the one hand, each agent has a higher incentive to approve A conditional on receiving signal h since this signal favors state θ_3 more. On the other hand, each agent has a lower incentive to approve A conditional on being pivotal since the distribution of the total number of approvals in state θ_2 moves closer to the distribution in state θ_1, which decreases this agent's posterior belief of state θ_3 conditional on being pivotal. Thus, a smaller ρ_2 has an ambiguous effect on the agents' incentives to approve A conditional on being pivotal and receiving signal h, and hence has an ambiguous effect on information transmission. We can apply a similar intuition to the case where ρ_3 increases. However, when ρ_1 decreases, the two effects mentioned above move the agents' posterior beliefs of state θ_3 in the same direction. Thus, a smaller ρ_1 increases the agents' incentives to approve A conditional on being pivotal and receiving signal h, and hence contributes to information transmission.

Corollary 1.3 indicates that the boundary of information transmission, that is, the red dashed line in Figure 1.7, moves outward when ρ_1 decreases. However, changes in ρ_3 and ρ_2 have ambiguous effects on it.

Corollary 1.4. *The threshold $\hat{\lambda}_2$ is not monotonic with λ_1.*

In Figure 1.8, we plot $\hat{\lambda}_{2,j}$ for $j \in \{1,2,3\}$ and $\hat{\lambda}_2$ as functions of λ_1. As stated before, an informative equilibrium with $\hat{T} = j$ exists if $\lambda_2 < \hat{\lambda}_{2,j}$ when N is large.

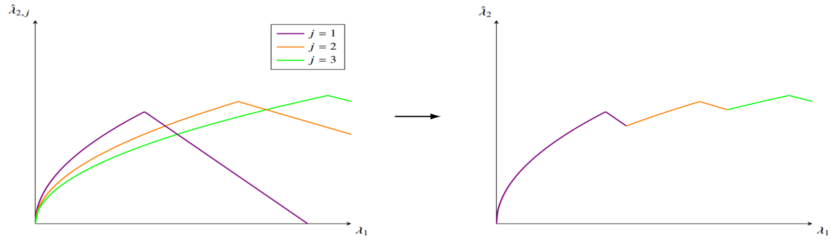

Figure 1.8. Non-monotonic boundaries

The threshold $\hat{\lambda}_{2,j}$ is not monotonic with λ_1 for each $j \in \mathbb{N}$. To see the intuition, let us fix an arbitrary $j \in \mathbb{N}^+$ and let the principal always choose $\hat{T} = j$. We then consider the case where λ_1 increases while other parameters are constant. The prior q_2^0 is smaller and hence the conflict of interest between the principal and the agents is smaller, which contributes to information transmission and increases $\hat{\lambda}_{2,j}$. However, the prior q_3^0 also decreases, and hence the state θ_3 is less likely, which decreases the agents' incentives to approve A when they receive signal h, which impedes information transmission and decreases $\hat{\lambda}_{2,j}$.

1.4.2 Ranking Informative Equilibria

When we fix all parameter values, there cannot exist more than one informative equilibrium in which the principal chooses the same $\hat{T}$ since there exists at most one x_h solving (1.6). However, there might exist multiple informative equilibria with different $\hat{T}$, as shown by the left panel of Figure 1.8. We now rank them according to the payoffs of the principal and the agents. Let $U_{pc}(\Gamma)$ and $U_{ag}(\Gamma)$ be the expected payoff of the principal and the expected payoff of the agents respectively for a given equilibrium Γ.

Proposition 1.6. *Fix all parameter values. If there exist two informative equilibria* $\Gamma_1 = \{x_{h,1}, \hat{T}_1\}$ *and* $\Gamma_2 = \{x_{h,2}, \hat{T}_2\}$ *such that* $\hat{T}_1 < \hat{T}_2$, *then*

$$\begin{aligned} x_{h,1} &\leq x_{h,2}, \\ U_{pc}(\Gamma_1) &\leq U_{pc}(\Gamma_2), \\ U_{ag}(\Gamma_1) &\leq U_{ag}(\Gamma_2). \end{aligned}$$

All inequalities are strict if $x_{h,1} < 1$.

When the principal requires a higher $\hat{T}$, the agents approve A with a higher probability and hence increase x_h. The principal observes N messages from the agents that are identically distributed and independent conditional on the state and makes his decision to maximize his expected payoff. When x_h is higher, each message is more Blackwell informative, and hence the joint N messages are also more Blackwell informative. Thus, the principal receives a higher expected payoff from the equilibrium with a higher $\hat{T}$.

For the agents, consider an informative equilibrium $\Gamma = \{x_h, \hat{T}\}$ with $x_h < 1$. Each agent is indifferent between A and B conditional on receiving signal h and being pivotal. Hence, she is indifferent conditional on the event that there are $\hat{T}$ approvals from N agents since the agents randomize when she receives signal h. Therefore, the principal would still choose $\hat{T}$ if he shared the same preference with the agents. Thus, an informative equilibrium with a higher $\hat{T}$ also benefits the agents since the principal chooses $\hat{T}$ under a more Blackwell informative information structure.

As discussed above, in every informative equilibrium, when we fix the strategy of the agents, the principal and the agents agree on the same threshold $\hat{T}$, that is, they share common interests. Therefore, we can rank all informative equilibria in the Blakweell order or the Pareto order, and these two orders coincide with each other.

From Proposition 1.6, the informative equilibrium with the highest cut-off $\hat{T}_{max}$ maximizes the expected payoffs of the principal and the agents among all informative equilibria. We denote this equilibrium by the ***most informative equilibrium***. Note that the agents also choose the highest x_h in the most informative equilibrium among all informative equilibria. From Proposition 1.3, the highest cut-off $\hat{T}_{max}$ is always smaller than a number that is independent of N since messages from the agents cannot be too informative. Otherwise, the inference from being pivotal overwhelms each agent's private information.

We can show that for almost all parameter values that satisfy $\frac{V_{ag}(\theta_1)}{V_{ag}(\theta_3)}\frac{V_{pc}(\theta_3)}{V_{pc}(\theta_1)} < \frac{\rho_3}{\rho_1}$ and $\lambda_2 < \hat{\lambda}_2$, that is, for almost all parameter values under which information transmission persists, the highest cut-off $\hat{T}_{max}$ is independent of N when N is above some threshold. For the other parameter values that satisfy both conditions above, the highest cut-off $\hat{T}_{max}$ takes a value between two adjacent numbers. In the left panel of Figure 1.8, we can see that when N is large, the cut-off $\hat{T}_{max} = 1$ if (λ_1, λ_2) is above the orange line and below the purple line while $\hat{T}_{max} = 2$ if (λ_1, λ_2) is above the green line and below the orange line. However, for some points of (λ_1, λ_2) exactly on the orange line, the highest cut-off $\hat{T}_{max}$ might be either 1 or 2 when N is large.

1.4.3 Amount of Information Transmission

In this section, we discuss the maximal amount of information transmission by focusing on the most informative equilibrium $\Gamma_{max} = \{x_{h,max}, \hat{T}_{max}\}$.

Proposition 1.7. *In the equilibrium Γ_{max}, the agents' equilibrium strategy $x_{h,max}$ increases with $V_{ag}(\theta_i)$ and decreases with $V_{pc}(\theta_i)$ for each $i \in \{1,2,3\}$. Furthermore, it decreases with λ_2.* [21]

As discussed in Section 1.4.2, the principal receives more information from the agents if the agents choose a higher x_h. By Proposition 1.7, the agents' equilibrium strategy $x_{h,max}$ decreases with $\frac{V_{ag}(\theta_1)}{V_{ag}(\theta_3)}\frac{V_{pc}(\theta_3)}{V_{pc}(\theta_1)}$ when we only vary one term, and also decreases with λ_2. Hence, the maximal amount of information transmission decreases with both types of conflict between the principal and the agents.

We now compare the maximal amount of information transmission as $N \to \infty$ in the setting with the misaligned state θ_2 with the one in the setting with no misaligned state analyzed by the existing literature. We measure the maximal amount of information transmission by[22]

$$I = \limsup_{N\to\infty} \frac{V_{pc}^{max} - V_{pc}^{0}}{V_{pc}^{full} - V_{pc}^{0}} \in [0,1],$$

where (i) V_{pc}^{max} is the principal's expected payoff from the most informative equilibrium, (ii) V_{pc}^{0} is the principal's expected payoff from the uninformative babbling equilibrium, and (iii) V_{pc}^{full} is the principal's expected payoff if he can observe the realized state.

Figure 1.9 illustrates the maximal amount of information transmission regarding the two types of conflict between the principal and the agents, the conflict generated by the different payoff intensities in state θ_1 and θ_3 that is measured by $\frac{V_{ag}(\theta_1)}{V_{ag}(\theta_3)}\frac{V_{pc}(\theta_3)}{V_{pc}(\theta_1)}$,[23] and the conflict concerning the misaligned state θ_2 that is measured by λ_2.

21. We always change one parameter and keep others including N fixed.
22. The limit always exists in the situations where $\lambda_2 = 0$.
23. When varying $\frac{V_{ag}(\theta_1)}{V_{ag}(\theta_3)}\frac{V_{pc}(\theta_3)}{V_{pc}(\theta_1)}$, we either only change $V_{ag}(\theta_1)$ or only change $V_{ag}(\theta_3)$, while keeping other parameters fixed.

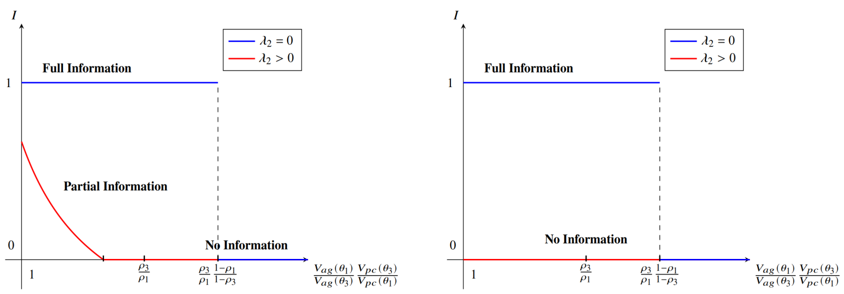

Figure 1.9. Maximal amount of information transmission

Notes: We plot I as a function of $\frac{V_{ag}(\theta_1)}{V_{ag}(\theta_3)}\frac{V_{pc}(\theta_3)}{V_{pc}(\theta_1)}$ in different cases. The blue line corresponds to the case where $\lambda_2 = 0$. The red line corresponds to the case where $\lambda_2 > 0$. In the left panel, we choose $\lambda_2 < \hat{\lambda}_2$ given that $\frac{V_{ag}(\theta_1)}{V_{ag}(\theta_3)}\frac{V_{pc}(\theta_3)}{V_{pc}(\theta_1)} = 1$ for the red line. In the right panel, we chooses $\lambda_2 > \hat{\lambda}_2$ given that $\frac{V_{ag}(\theta_1)}{V_{ag}(\theta_3)}\frac{V_{pc}(\theta_3)}{V_{pc}(\theta_1)} = 1$ for the red line.

When $\lambda_2 = 0$, Proposition 1.1 shows that the principal fully learns the state as $N \to \infty$ if $\frac{V_{ag}(\theta_1)}{V_{ag}(\theta_3)}\frac{V_{pc}(\theta_3)}{V_{pc}(\theta_1)}$ is below the threshold $\frac{\rho_3}{\rho_1}\frac{1-\rho_1}{1-\rho_3}$. Otherwise, information transmission fails and only the babbling equilibrium exists.

This paper analyzes the setting with $\lambda_2 > 0$. In the left panel with a small λ_2, even if the principal and the agents have the same payoffs in state θ_1 and state θ_3 with

$$\frac{V_{ag}(\theta_1)}{V_{ag}(\theta_3)} \cdot \frac{V_{pc}(\theta_3)}{V_{pc}(\theta_1)} = 1,$$

information aggregation fails and the amount of information transmission is limited. As $\frac{V_{ag}(\theta_1)}{V_{ag}(\theta_3)}\frac{V_{pc}(\theta_3)}{V_{pc}(\theta_1)}$ increases, the principal receives less information. When $\frac{V_{ag}(\theta_1)}{V_{ag}(\theta_3)}\frac{V_{pc}(\theta_3)}{V_{pc}(\theta_1)}$ is above a threshold that is lower than $\frac{\rho_3}{\rho_1}$, the principal receives no information. In this case, information transmission fails and only the babbling equilibrium exists. Note that the threshold above which information transmission fails decreases with λ_2. In the right panel with a large λ_2, information transmission always fails according to Proposition 1.5.

1.4.4 Information Aggregation and Transmission

We claim that the failure of information aggregation might facilitate information transmission and further argue that communication among the agents might impede information transmission and hurt both the principal and the agents.

Consider the case where there is only one agent and this agent receives all N signals. She advises the principal to choose A or not. As $N \to \infty$, this agent is fully informed about the realized state. She advises the principal to choose A in state θ_3 and choose B in state θ_2 and state θ_1. The principal follows this agent's advice of choosing B if he receives a negative expected payoff from choosing A,

$$q_2^0 V_{pc}(\theta_2) + q_1^0 V_{pc}(\theta_1) < 0,$$

that is, if

$$\frac{\lambda_2}{\lambda_1} < -\frac{V_{pc}(\theta_1)}{V_{pc}(\theta_2)}.$$

Proposition 1.8. *There exists $\bar{\lambda}_1$ such that*[24]

$$\frac{\hat{\lambda}_2}{\lambda_1} > -\frac{V_{pc}(\theta_1)}{V_{pc}(\theta_2)} \quad \textit{iff} \quad \lambda_1 < \bar{\lambda}_1.$$

We plot $\hat{\lambda}_2$ as a function of λ_1 in Figure 1.10, which illustrates Proposition 1.8. Consider a pair of (λ_1, λ_2) in the shaded area. If there are N agents and each of them receives a private signal, full information aggregation fails but information transmission persists since $\lambda_2 < \hat{\lambda}_2$. If there is only one agent who receives all N signals, as $N \to \infty$, she fully learns the realized state but information transmission fails since

$$\lambda_2 V_{pc}(\theta_2) + \lambda_1 V_{pc}(\theta_1) > 0.$$

The principal chooses A even if the agent advises him to choose B.

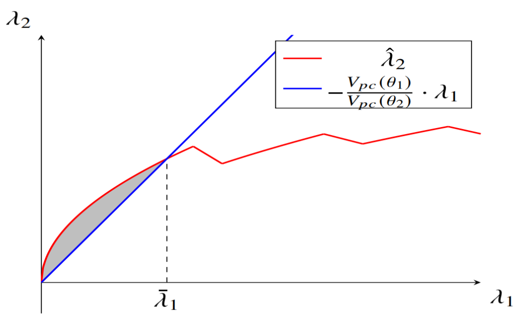

Figure 1.10. Information transmission with the failure of information aggregation

The intuition for the argument that the failure of information aggregation might facilitate information transmission goes as follows. Many studies in cheap-talk literature, initiated by Crawford and Sobel (1982) consider a model of information transmission between one sender and one receiver. They show that the sender might make her message intentionally vague since intentional vagueness mitigates the conflict of interest between the sender and the receiver and further facilitates information transmission. Now, we can also interpret the failure of information aggregation as intentional vagueness if we regard all N agents as the sender and the principal as the receiver. Such intentional vagueness disappears

24. The value of $\bar{\lambda}_1$ depends on the value of other parameters except λ_1 and N.

in the case where an agent fully learns the state but does not have commitment power.

Furthermore, we can show that for each (λ_1, λ_2) in the shaded area, there always exists an informative equilibrium with $\hat{T} = 1$ when N is large since the unanimity rule aggregates information the least efficiently and hence generates the largest intentional vagueness.

Both the principal and the agents benefit from the failure of information aggregation when (λ_1, λ_2) is in the shaded area since both of them receive higher payoffs from any informative equilibrium than from the babbling equilibrium. Hence, it might be better to disperse the information among the agents instead of letting an agent receive all the signals when this agent cannot commit to generating intentional vagueness.

An important finding of Battaglini (2017) is that communication among the agents facilitates information transmission and aggregation, benefiting both the principal and the agents. He hence highlights the value of social media to the effectiveness of petitions and public protests, since social media allow citizens to share information. In contrast, we show that the communication among the agents might impede information transmission and hurt both the principal and the agents. Note that the case where the agents fully communicate with each other and share their signals is equivalent to the case analyzed above in which there is only one agent and this agent receives all N signals.

1.5 Beyond the Binary Situation

We now extend the model to the case where neither the signal space nor the message space is binary. Each agent $i \in \{1, \ldots, N\}$ receives a private signal $s^i \in \{s_1, \ldots, s_J\}$ with $J \geq 2$. The signals are identically distributed and independent across the agents conditional on the state $\theta \in \{\theta_1, \theta_2, \theta_3\}$. There exists $\alpha > 0$ such that

$$\mathbb{P}[s_j|\theta] > \alpha, \ \forall j \in \{1, \ldots, J\} \text{ and } \theta \in \{\theta_1, \theta_2, \theta_3\}.$$

That is, an agent cannot exclude any state if she receives a particular signal. We generalize (1.3) by assuming the strict ***Monotone Likelihood Ratio Property (MLRP)***:

$$\frac{\mathbb{P}[s_j|\theta_3]}{\mathbb{P}[s_j|\theta_2]} \text{ and } \frac{\mathbb{P}[s_j|\theta_2]}{\mathbb{P}[s_j|\theta_1]} \text{ strictly increase with } j. \tag{1.12}$$

Each agent i can send a message $z^i \in \{z_1, \ldots, z_K\}$ with $K \geq 2$. The principal observes $T = (T_1, \ldots, T_K) \in \Delta^K(N)$, that is, the total number of each kind of message, and chooses between A and B.

In the example of non-binding shareholder voting, besides voting in favor of or rejecting the new proposal, the shareholders can also stay neutral and abstain. Similarly, in the example of public protests, the citizens can choose among joining the rally for implementing reform, joining the rally for keeping the *status-quo*, or staying neutral and remaining silent. We can also use this framework to study survey polls in which each interviewee sends a score rating the desirability of a new policy.

We examine symmetric Bayesian Nash equilibrium in which the agents use the same strategy $P = \{p_{j,k}\}_{J\times K}$ such that an agent sends the message z_k with a probability $p_{j,k}$ when she receives the signal s_j. The strategy of the principal is a function

$$\psi : \Delta^K(N) \rightarrow [0,1]$$

such that he chooses A with probability $\psi(T)$ when he observes $T = (T_1, ..., T_K)$.

The agents follow a ***monotonic strategy*** if they are more likely to send higher messages when they receive higher signals, that is,[25]

$$p_{j',k} \cdot p_{j,k'} \leq p_{j,k} \cdot p_{j',k'} \quad \text{for each } j < j' \text{ and } k < k'. \tag{1.13}$$

The principal follows a ***monotonic strategy*** if he chooses A with a higher probability when an agent switches from a lower message to a higher one, that is, for each $T = (T_1, ..., T_K) \in \Delta^K(N-1)$ and each $m < m'$,

$$\psi(T_1, ..., T_m + 1, ..., T_{m'}, ..., T_K) \leq \psi(T_1, ..., T_m, ..., T_{m'} + 1, ..., T_K).$$

Note that when one side uses a monotonic strategy, it is without loss of generality to let the other side use a monotonic strategy as the best response. We focus on the ***monotonic equilibrium*** in which both the principal and the agents use monotonic strategies.

Monotonic equilibria are reasonable and fit applications well while non-monotonic equilibria are counterintuitive and hard to be implemented. Intuitively, a shareholder should support the new proposal more, a citizen should be more likely to quit the rally for keeping the *status-quo* and join the one for implementing reform, and an interviewee should rate the new policy with a higher score if they are more optimistic about the new proposal, reform, or new policy based on their private information. It is also reasonable that a manager should accept the new proposal with a higher probability if fewer shareholders object to it or more shareholders support it, a politician should implement reform with a higher probability if fewer citizens join in the rally for keeping the *status-quo* or more citizens join the rally for implementing the reform, and an interviewer should choose the new policy with a higher probability if more interviewees rate it with higher scores. There is growling literature studying the monotonic equilibrium in communication games, as discussed in Section 1.7.

25. It is equivalent to $\frac{p_{j',k}}{p_{j,k}} \leq \frac{p_{j',k'}}{p_{j,k'}}$ when both $p_{j,k}$ and $p_{j,k'}$ are positive.

Proposition 1.9. *For each $\epsilon > 0$, there exists N''_ϵ such that for each $N > N''_\epsilon$, in every monotonic equilibrium except the babbling one, the agents only send z_1[26] when they receive $s \in \{s_1, ..., s_{J-1}\}$ and send z_1 with probability larger than $1-\epsilon$ when they receive s_J, that is,*

$$\begin{aligned} p_{j,1} &= 1, \qquad \forall\, j \in \{1, ..., J-1\}, \\ p_{J,1} &> 1-\epsilon. \end{aligned}$$

By Proposition 1.9, when N is large enough, there is no difference among $\{z_2, ..., z_K\}$. The agents send these messages only if $s = s_J$. Therefore, when N is large enough, we return to the basic model with binary signals and binary messages such that

$$\rho_i = \mathbb{P}[s_J | \theta_i], \ \forall i \in \{1, 2, 3\}.$$

Hence, we can easily extend all results in Section 1.3.2 and Section 1.4. In particular,

Theorem 1.2. *No sequence of monotonic equilibria aggregates information. That is, there exists a constant $c > 0$ such that for each N and each monotonic equilibrium Γ with N agents,*

$$\mathbb{P}[A|\theta_1; \Gamma] + \mathbb{P}[B|\theta_3; \Gamma] > c.$$

We now sketch the proof for Proposition 1.9. For simplicity, consider the case where each agent can send a message $z \in \{z_1, z_2, z_3\}$. Let the agents use a monotonic strategy P. From (1.12), (1.13) and some additional regularity assumptions concerning P to avoid degenerate cases, we can show that the distributions of the message from an agent also satisfy strict MLRP,

$$\frac{\mathbb{P}[z_k|\theta_3]}{\mathbb{P}[z_k|\theta_2]} \text{ and } \frac{\mathbb{P}[z_k|\theta_2]}{\mathbb{P}[z_k|\theta_1]} \text{ strictly increase with } k.$$

Denote the distributions of the message from an agent in θ_1, θ_2, and θ_3 by G_1, G_2, and G_3 respectively. We have

$$G_3 \succ G_2 \succ G_1 \tag{1.14}$$

in the monotone likelihood ratio order.

Consider the ***set of pivotal events***,

$$E_N = \{T = (T_1, T_2, T_3) \in \Delta^3(N-1) \mid \psi(T_1+1, T_2, T_3) \neq \psi(T_1, T_2, T_3+1)\},$$

26. We ignore the degenerate case where agents never send z_1. In this case, just relabel the lowest message that the agents send with a positive probability by z_1.

that is, consider the set of events in which one additional message might change the principal's decision.[27]

Now, let us fix the strategy of the agents and let $N \to \infty$. We can show that

$$\lim_{N\to\infty} \frac{\mathbb{P}[E_N|\theta_3]}{\mathbb{P}[E_N|\theta_2]} = 0 \text{ and } \lim_{N\to\infty} \frac{\mathbb{P}[E_N|\theta_3]}{\mathbb{P}[E_N|\theta_1]} = 0. \tag{1.15}$$

Intuitively, given that the principal is not sure whether the realized state is θ_1 or not, the realized state must be either θ_1 or θ_2 since the "distance" between distributions G_1 and G_2 is smaller than the one between G_1 and G_3 according to (1.14). We extend the intuition in Section 1.3.2 to higher dimensions.

To see more precisely why (1.15) is true, consider the posterior likelihood ratio for $T \in \Delta^3(N-1)$ and each $i, i' \in \{1,2,3\}$,

$$\begin{aligned} \frac{\mathbb{P}[T|\theta_i]}{\mathbb{P}[T|\theta_{i'}]} &= \prod_{k=1}^{3}\left[\frac{\mathbb{P}[z_k|\theta_i]}{\mathbb{P}[z_k|\theta_{i'}]}\right]^{T_k} = \exp\left\{\sum_{k=1}^{3} T_k \log\frac{\mathbb{P}[z_k|\theta_i]}{\mathbb{P}[z_k|\theta_{i'}]}\right\} \\ &= \exp\left\{(N-1)\cdot\left[KL\left(\gamma(T),G_{i'}\right) - KL\left(\gamma(T),G_i\right)\right]\right\}, \end{aligned} \tag{1.16}$$

where $\gamma(T)$ is the sample frequency with

$$\gamma(T) = \left(\gamma_1(T), \gamma_2(T), \gamma_3(T)\right) = \left(\frac{T_1}{N-1}, \frac{T_2}{N-1}, \frac{T_3}{N-1}\right),$$

and $KL(\cdot,\cdot)$ is the Kullback–Leibler divergence (KL divergence) with

$$KL(\gamma, G_i) = \sum_{k=1}^{3} \gamma_k \log\frac{\gamma_k}{\mathbb{P}[z_k|\theta_i]}, \quad \forall i \in \{1,2,3\}.$$

It measures how γ (observed frequency) deviates from G_i (mean in state θ_i). The larger $KL(\gamma, G_i)$ is, the more rare that a sample with a frequency γ in state θ_i is.

From (1.16), as when N is large, instead of focusing on the set of pivotal events E_N, we can work with the ***set of pivotal frequencies***,

$$F = \left\{\gamma \in \Delta^3(1) | KL(\gamma, G_1) = \min\left[KL(\gamma, G_2), KL(\gamma, G_3)\right]\right\}.$$

For each $\gamma \notin F$, we have $T \notin E_N$ for each T with $\gamma(T) = \gamma$ when N is large. For example, consider a $\tilde{\gamma} \notin F$ such that

$$KL(\tilde{\gamma}, G_1) < \min\left[KL(\tilde{\gamma}, G_2), KL(\tilde{\gamma}, G_3)\right].$$

When N is large and the principal observes $\tilde{T}$ from $N-1$ agents such that $\gamma(\tilde{T}) = \tilde{\gamma}$, he must be sure that the state is θ_1 by (1.16). Hence, one additional message cannot change his decision. We have $\tilde{T} \notin E_N$.

27. Note the in any monotonic equilibrium except the babbling equilibrium, for each $T = (T_1, T_2, T_3) \in \Delta^3(N-1)$, if $\psi(T_1+1, T_2, T_3) \neq \psi(T_1, T_2+1, T_3)$ or $\psi(T_1, T_2+1, T_3) \neq \psi(T_1, T_2, T_3+1)$, then we must have $\psi(T_1+1, T_2, T_3) \neq \psi(T_1, T_2, T_3+1)$.

Note that in the binary setting analyzed in Section 1.3.2, the set F is a singleton, which is not true when we move beyond the binary setting. We provide a way to identify the unique most likely pivotal frequency

$$\gamma^* = \arg\min_{\gamma \in F} KL(\gamma, G_1).$$

Note that we only need to consider the pivotal events with frequencies concentrated around γ^* since the unconditional likelihoods of them dominate the unconditional likelihoods of other pivotal events at an exponential rate as shown in (1.16). We further show that

$$KL(\gamma^*, G_1) = KL(\gamma^*, G_2) < KL(\gamma^*, G_3). \tag{1.17}$$

We prove (1.15) by using (1.16) and (1.17).

To find the most likely pivotal frequency, let us consider a type of statistical distance between distribution G_i and $G_{i'}$ for $i \neq i'$, the ***Chernoff Information***:

$$c(G_i, G_{i'}) = \min_{\gamma \in \Delta^3(1)} KL(\gamma, G_i) \quad s.t.\ KL(\gamma, G_i) = KL(\gamma, G_{i'}).$$

The minimizing problem has a unique minimizer.[28] Denote it by $\gamma_{i,i'}$ or $\gamma_{i',i}$.

It can be show that[29]

$$c(G_1, G_2) < c(G_1, G_3),$$

if (1.14) is satisfied. We further show that

$$KL(\gamma_{1,2}, G_1) = KL(\gamma_{1,2}, G_2) < KL(\gamma_{1,2}, G_3).$$

The key step is to show that

$$G_3 \succ G_2 \succ \gamma_{1,2} \succ G_1$$

in the monotonic likelihood ratio order if we regard $\gamma_{1,2}$ as a signal distribution. Therefore, the frequency $\gamma_{1,2}$ is the most likely pivotal frequency and satisfies (1.17).

We plot the simplex of distributions and frequencies in Figure 1.11, which illustrates the reasoning presented above.

28. Both the function $KL(\gamma, G_i)$ and the set $\{\gamma \in \Delta^3(1) \mid KL(\gamma, G_i) = KL(\gamma, G_{i'})\}$ are convex.
29. Frick, Iijima, and Ishii (2021a) first find this result. There will be a note forthcoming for further discussion.

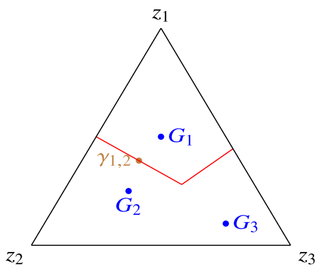

Figure 1.11. The most likely pivotal frequency

Notes: The corner z_k corresponds to the distribution or the sample frequency that the message is always z_k. The segment $z_k z_{k'}$ corresponds to the set of distributions and sample frequencies that the message is always either z_k or $z_{k'}$. The red line is the set of pivotal frequencies from which the "distance" (KL divergence) to G_1 equals the minimum of distances to G_2 and G_3. Among all pivotal frequencies, the frequency $\gamma_{1,2}$ has the shortest distance to G_1 and hence it is the most likely pivotal frequency.

We show that in every sequence of monotonic equilibria, we must have

$$\lim_{N\to\infty} KL(G_i, G_{i'}) = 0, \quad \forall i, i' \in \{1,2,3\},$$

$$\lim_{N\to\infty} c(G_i, G_{i'}) = 0, \quad \forall i, i' \in \{1,2,3\},$$

that is, the distributions of messages in different states must be close to each other. Otherwise, the agents realize that the state must be either θ_1 or θ_2 conditional on being pivotal by (1.15) and ignore their own signals. Hence, the information contained in an agent's message must vanish as $N \to \infty$.

By extending Lemma 1.1, we show that the agents only send the lowest message z_1 when they receive the lowest signal s_1. Finally, we demonstrate that as N grows large, the agents only send z_1 when they receive $s \in \{s_1, \ldots, s_{J-1}\}$ and send z_1 with probability near 1 when they receive $s = s_J$ since (i) the strategy of the agents must be a monotonic mapping and satisfy a single crossing condition[30] due to the strict MLRP of signals, and (ii) information contained in an agent's message must vanish.

1.6 Commitment Case

In the basic model, the principal cannot *ex-ante* commit to a decision rule. Consequently, conditional on being pivotal, each agent learns that the principal must be nearly indifferent between A and B and infers the realized state from such an event. We now consider the case where the principal can design and commit to

30. That is, for each $k' > k$, if the agents send message z_k with a positive probability when they receive signal s_j, then they never send message $z_{k'}$ when they receive signals $s_{j'}$ with $j' < j$.

a decision mechanism. In this section, we show that the principal can approach his first-best outcome as $N \to \infty$ by committing to mechanisms with a simple structure.

With no misaligned state

We first consider the case with no misaligned state, that is, the case where $q_2^0 = 0$. Let us start with direct and anonymous mechanisms that depend only on T, that is, the total number of agents reporting signal h. The principal commits to a ***cut-off mechanism*** if there exists $\hat{T} \in \mathbb{N}$ such that the principal chooses A when $T \geq \hat{T}$ and choose B otherwise.

Note that when the agents can observe all signals together, for each N, there exists a cut-off $\bar{T}_N$ such that the agents prefer A if and only if more than $\bar{T}_N$ of them receive signal h. The principal then commits to a sequence of cut-off mechanisms $\{\hat{T}_N\}_{N=1}^{\infty}$ with $\hat{T}_N = \bar{T}_N$ for each N. It is always incentive compatible for the agents to report truthfully. The principal can approach his first-best outcome as $N \to \infty$.

The principal can also pick any $t \in (0,1)$ and run an election among the agents following a qualified majority rule with t, in which the agents choose whether to vote for A, and A is chosen if the ratio of votes for it exceeds t. By the Condorcet jury theorem and its modern versions (Feddersen and Pesendorfer (1997), Feddersen and Pesendorfer (1998),Myerson (1998), Duggan and Martinelli (2001)), as $N \to \infty$, dispersed information among the agents is effectively aggregated and the principal approaches his first-best outcome.

With the misaligned state

We now consider the case with the misaligned state θ_2, that is, with $q_2^0 > 0$. First, the principal cannot approach his first-best outcome by committing to a sequence of cut-off mechanisms. Figure 1.12 illustrates this argument. When N is large, the principal must choose $\hat{T}_N \in (N\rho_1, N\rho_2)$ to approach his first-best outcome. However, each agent realizes that the state must be either θ_1 or θ_2 conditional on being pivotal, that is, conditional on the event that from the other $N-1$ agents, $\hat{T}_N - 1$ of them receive signal h. She does not have the incentive to report truthfully when she receives signal h.

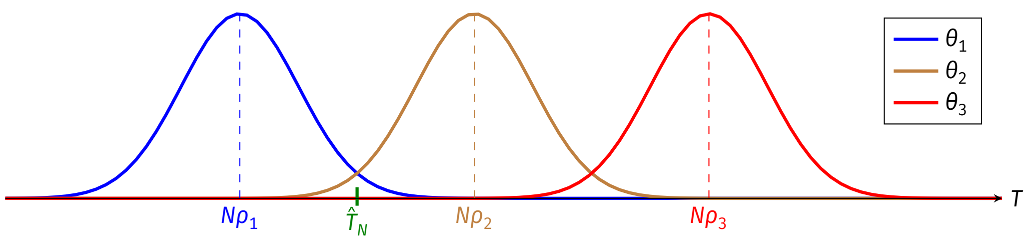

Figure 1.12. Distributions of the total number of signal h

Notes: Distributions of the total number of signal h and one cut-off $\hat{T}_N$. The cut-off mechanism with $\frac{\hat{T}_N}{N} \in (\rho_1, \rho_2)$ is not incentive compatible.

Furthermore, for each $t \in (0,1)$, the principal cannot approach his first-best outcome by committing to an election following a qualified majority rule with t, in which according to the Condorcet jury theorem, information is aggregated as $N \to \infty$ but the agents approach their first-best outcome.

The principal can approach his first-best outcome by mixing two cut-off mechanisms. Consider a mechanism $M = (\mu, \hat{T}^{\alpha}, \hat{T}^{\beta})$ such that the principal commits to choosing the cut-off mechanism $\hat{T}^{\alpha}$ with probability μ and choosing the cut-off mechanism $\hat{T}^{\beta}$ with probability $1-\mu$. The agents cannot observe the principal's choice.

Proposition 1.10. *there exists a sequence of mechanisms $\{M_N\}_{N=1}^{\infty}$ with $M_N = (\mu_N, \hat{T}_N^{\alpha}, \hat{T}_N^{\beta},)$ for each N such that*

$$\begin{aligned}
\lim_{N\to\infty} Pr(A|\ \theta_3; M_N) &= 1,\\
\lim_{N\to\infty} Pr(A|\ \theta_2; M_N) &= 1,\\
\lim_{N\to\infty} Pr(B|\ \theta_1; M_N) &= 1.
\end{aligned}$$

Note that each agent makes a decision conditional on being pivotal, that is, conditional on the event that her report can change the decision of the principal. In the mechanism $M_N = (\mu_N, \hat{T}_N^{\alpha}, \hat{T}_N^{\beta},)$ with N agents, when the principal chooses $\hat{T}_N^{\alpha}$, an agent is pivotal if from the other $N-1$ agents, there are $\hat{T}_N^{\alpha}-1$ agents who receive signal h, and when the principal chooses $\hat{T}_N^{\beta}$, an agent is pivotal if from the other $N-1$ agents, there are $\hat{T}_N^{\beta}-1$ agents who receive signal h. The principal hence can mix between $\hat{T}_N^{\alpha}$ and $\hat{T}_N^{\beta}$ to manipulate the agents' inferences from being pivotal.

As illustrated by Figure 1.13, the principal chooses $\hat{T}_N^{\alpha} \in (N\rho_1, N\rho_2)$ for his first-best outcome and chooses $\hat{T}_N^{\beta}$ close to $N\rho_3$ to fulfill the incentive-compatible constraint. Note that the agents prefer B conditional on being pivotal in the cut-off mechanism with $\hat{T}_N^{\alpha}$, while they prefer A conditional on being pivotal in the cut-off mechanism with $\hat{T}_N^{\beta}$. By mixing between $\hat{T}_N^{\alpha}$ and $\hat{T}_N^{\beta}$, the principal can make the agents indifferent between A and B conditional on being pivotal. Hence, they have incentives to report their signals truthfully. Furthermore, by choosing $\hat{T}_N^{\beta}$ close to $N\rho_3$, the principal can choose $\hat{T}_N^{\beta}$ with a probability approaching 0 as $N \to \infty$. He pays almost no information rent to the agents and approaches his first-best outcome.

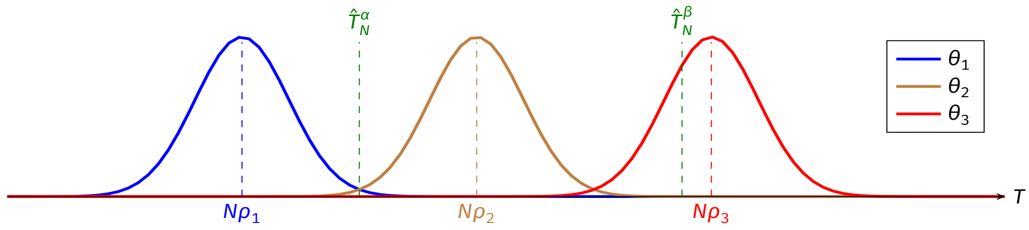

Figure 1.13. Distributions of the total number of signal h

Notes: Distributions of the total number of signal h and two cut-offs $\hat{T}_N^{\alpha}$ and $\hat{T}_N^{\beta}$. The principal mixes between two cut-offs and lets $\frac{\hat{T}_N^{\beta}}{N}$ be close to ρ_3.

Similarly, the principal can approach his first-best outcome by randomizing between two qualified majority rules with different t.

1.7 Related Literature

This paper is related to the literature on cheap talk with multiple senders. This paper further considers the case where senders (agents) have the same preference. Besides Levit and Malenko (2011) and Battaglini (2017), Wolinsky (2002) analyzes a similar model and also shows that information transmission fails and complete unraveling happens if the conflict of interest between the principal and the agents is large. Ekmekci and Lauermann (2022) follow the setting of Battaglini (2017) but add costly participation, that is, each agent in our basic model needs to pay a cost drawn from a distribution when rejecting A.[31] They show that information is aggregated even if the conflict of interest is above the threshold given by Levit and Malenko (2011) and Battaglini (2017). However, It is ambiguous whether a similar result holds in our setting with the misaligned state θ_2. Morgan and Stocken (2008) consider a model where the agents have heterogeneous preferences. They show that when the principal and the agents have similar preferences, information is effectively aggregated and the principal fully learns the state when the number of agents grows large.

This paper is also related to the literature on information aggregation in elections (Feddersen and Pesendorfer (1997), Feddersen and Pesendorfer (1998), Myerson (1998), Duggan and Martinelli (2001)), which demonstrates that information dispersed among voters is effectively aggregated in elections with predetermined qualified majority rules while information aggregation fails under the unanimity rule. In contrast, we consider the case where the principal cannot *ex-ante* commit to a rule. We show that he optimally follows the unanimity rule or rules close to it, and information aggregation fails. Razin (2003) considers a novel model in which the voters vote between two candidates and the winning

31. In this case, approving A is the default choice for the agents.

candidate chooses the policy based on his own decision. He shows that if both candidates have large conflicts of interest with voters, full information aggregation fails in a special subset of symmetric equilibria under a symmetric setting. We consider a different setting and show that the principal always fails to fully learn the state in all symmetric equilibria, even if the conflict of interest between him and the agents is small.

We further demonstrate that if the principal has the commitment power, he cannot approach his first-best outcome by committing to a qualified majority rule.[32] However, the principal can approach his first-best outcome by committing to randomizing between two qualified majority rules to manipulate the agents' inferences from being pivotal. Gerardi, McLean, and Postlewaite (2009) consider a similar mechanism in which the principal mixes between asking different numbers of agents to manipulate the agents' inferences from being pivotal.

We take inspiration from the literature on comparisons of statistical experiments. Moscarini and Smith (2002) consider a model in which a decision-maker is uncertain about the state of the world but can draw signals that are identically distributed and independent conditional on the state by performing an experiment repeatedly. Frick, Iijima, and Ishii (2021b) further consider the case of misspecified learning. They both provide rankings over statistical experiments by calculating the expected payoff of the decision-maker when he can perform a large number of experiments. Both rankings depend critically on Chernoff's information introduced in Section 1.6. Note that when the number of experiments approaches infinity, the speed at which the belief of the decision-maker converges depends crucially on the most likely events in which the principal stays uncertain about the realized state and hence the frequency that we derive when calculating Chernoff's information.

Finally, Section 1.6 contributes to the growing literature on the monotonic equilibrium in communication games, including Cho and Sobel (1990), Krishna and Morgan (2001), Chen, Kartik, and Sobel (2008), Ivanov (2010), Gordon et al. (2021), Kolotilin and Li (2021), and Vida, Honryo, and Azacis (2022). In addition, most of the literature studying the communication game in which each sender receives a noisy signal about the unknown state, including Austen-Smith (1990), Austen-Smith (1993), Morgan and Stocken (2008), Hagenbach and Koessler (2010), Galeotti, Ghiglino, and Squintani (2013)), and Currarini, Ursino, and Chand (2020) among others, focuses on binary signals and messages like our basic model, which guarantees the monotonicity of the equilibrium. The proof of Proposition 1.9 provides a novel and tractable way to analyze the case with multiple signals and messages.

32. However, the voters might obtain information aggregation, as discussed in Section 1.4.4.

1.8 Concluding Remarks

This paper analyzes a model of learning from multiple agents. In contrast to the existing literature, this paper considers the situation in which the preferences of the principal and the agents might not be completely aligned even if they fully know the state of the world, and introduce a different way to model the conflict of interest between the principal and the agents. The paper provides new insights regarding information transmission and demonstrates that learning is always incomplete no matter how many agents there are.

One promising direction for future research is to understand the effect of the number of agents on the welfare of the principal and the agents. We can show that in some situations, the expected payoffs of both the principal and the agents are maximized if the number of agents equals some finite number, while both the principal and the agents receive a lower expected payoff when we let the number of agents go to infinity, whenever we focus on the sequence of informative equilibria that maximize the welfare of the principal and the agents or the sequence of informative equilibria that minimize the welfare. As discussed in Section 1.4.3, there exist situations in which informative equilibria exist only if the number of agents is below some threshold. Furthermore, even when information transmission persists, we can find situations in which the maximal amount of information transmission is non-monotonic with the number of agents, and more surprisingly, the maximal amount of information transmission, the cut-off chosen by the principal, and the expected payoffs of the principal and the agents, are all maximized when the number of agents equals to a finite number. We expect that such results should hold generally for all parameter values.

Appendix 1.A Proofs

The appendices proceed as follows:

(1) In Appendix 1.A.1, we prove Lemma 1.1, Proposition 1.2, Proposition 1.3.

(2) In Appendix 1.A.2, we provide sufficient and necessary conditions under which there exists an informative equilibrium with $\hat{T} = 1$, that is, the equilibrium in which the principal chooses the unanimity rule. We then extend this approach for the informative equilibrium with $\hat{T} = i$ for each $i \in \mathbb{N}$.

(3) In Appendix 1.A.3, we prove the results in Section (1.4) based on results in Appendix 1.A.2.

(4) In Appendix 1.A.4, we characterize the equilibria in which the principal uses mixed strategies and demonstrate that it is without loss of generality to focus on the equilibria in which the principal uses pure strategies.

(5) In Appendix 1.A.5, we construct the mechanisms in which the principal approaches his first-best outcome as $N \to \infty$.

1.A.1 Proof of Lemma 1.1, Proposition 1.2, Proposition 1.3

1.A.1.1 Proof of Lemma 1.1

Assume there exists an informative equilibrium with $x_\ell > 0$. From (1.6) and (1.7), we have $x_h = 1$. Hence, we have $x_\ell \in (0,1)$. From (1.6),

$$\frac{q_3^0 \cdot (1-\rho_3) \cdot \mathbb{P}[piv|\theta_3; \boldsymbol{x}, \hat{T}] \cdot V_{ag}(\theta_3)}{-\sum_{i=1}^2 q_i^0 \cdot (1-\rho_i) \cdot \mathbb{P}[piv|\theta_i; \boldsymbol{x}, \hat{T}] \cdot V_{ag}(\theta_i)} = 1. \tag{1.A.1}$$

In state $\theta_i \in \{\theta_1, \theta_2, \theta_3\}$, the probability that one agent rejects A is $(1-\rho_i)(1-x_\ell)$. Hence,

$$\frac{(1-\rho_i) \cdot \mathbb{P}[piv|\theta_i; \boldsymbol{x}, \hat{T}]}{(1-\rho_{i'}) \cdot \mathbb{P}[piv|\theta_{i'}; \boldsymbol{x}, \hat{T}]} = \frac{(1-\rho_i) \cdot \mathbb{P}[T-1; N-1|\theta_i]}{(1-\rho_{i'}) \cdot \mathbb{P}[T-1; N-1|\theta_{i'}]} = \frac{\mathbb{P}[T-1; N-1|\theta_i]}{\mathbb{P}[T-1; N-1|\theta_{i'}]}.$$

Plug it into (1.A.1),

$$\frac{q_3^0 \cdot \mathbb{P}[T-1; N|\theta_3] \cdot V_{ag}(\theta_3)}{-\sum_{i=1}^2 q_i^0 \cdot \mathbb{P}[T-1; N|\theta_i] \cdot V_{ag}(\theta_i)} = 1.$$

From (1.1) and (1.4),

$$\frac{\sum_{i=2}^3 q_i^0 \cdot \mathbb{P}[T-1; N|\theta_i] \cdot V_{pc}(\theta_i)}{-q_1^0 \cdot \mathbb{P}[T-1; N|\theta_1] \cdot V_{pc}(\theta_1)} > 1.$$

That is,

$$L_{pc}(\hat{T}-1; \boldsymbol{x}) > 1,$$

which contradicts (1.8).

1.A.1.2 Proof of Proposition 1.2

Consider a strategy profile that the agents choose $x_h \in (0,1)$ and $x_\ell = 0$, we have

$$\begin{aligned}\frac{\mathbb{P}[T; N|\theta_i]}{\mathbb{P}[T; N|\theta_{i'}]} &= \frac{(\rho_i x_h)^T (1-\rho_i x_h)^{N-T}}{(\rho_{i'} x_h)^T (1-\rho_{i'} x_h)^{N-T}} = \exp\left[T \cdot \log \frac{\rho_i x_h}{\rho_{i'} x_h} + (N-T) \cdot \log \frac{1-\rho_i x_h}{1-\rho_{i'} x_h}\right] \\ &= \exp\left\{N \cdot \left[KL\left(\frac{T}{N}, \rho_{i'} x_h\right) - KL\left(\frac{T}{N}, \rho_i x_h\right)\right]\right\},\end{aligned} \tag{1.A.2}$$

where $KL(\cdot,\cdot)$ is the relative entropy with

$$KL(x,y) = x \log \frac{x}{y} + (1-x) \log \frac{1-x}{1-y}.$$

Fix some arbitrary $x \in (0,1)$. Consider a sequence of informative equilibrium $\{\Gamma_N = (x_{h,N}, \hat{T}_N)\}$ with

$$\lim_{N\to\infty} x_{h,N} = x. \tag{1.A.3}$$

We first claim that there exists $\hat{N}_1$ such that when $N > \hat{N}_1$, we must have

$$\hat{T}_N < N \cdot \rho_2 x_{h,N}. \tag{1.A.4}$$

Note that if (1.A.4) does not hold,

$$\begin{aligned} KL(\frac{\hat{T}_N}{N}, \rho_1 x_{h,N}) - KL(\frac{\hat{T}_N}{N}, \rho_2 x_{h,N}) &= \frac{\hat{T}_N}{N} \log \frac{\rho_2 x_{h,N}}{\rho_1 x_{h,N}} + (1 - \frac{\hat{T}_N}{N}) \log \frac{1-\rho_2 x_{h,N}}{1-\rho_1 x_{h,N}} \\ &\geq \rho_2 x_{h,N} \cdot \log \frac{\rho_2 x_{h,N}}{\rho_1 x_{h,N}} + (1-\rho_2 x_{h,N}) \log \frac{1-\rho_2 x_{h,N}}{1-\rho_1 x_{h,N}} \\ &= KL(\rho_2 x_{h,N}, \rho_1 x_{h,N}) \\ &> 0. \end{aligned}$$

The first inequality is from taking the derivative in $\frac{\hat{T}_N}{N}$. The second inequality is a result known as Gibbs' inequality. From (1.A.3), we can see that $KL(\rho_2 x_{h,N}, \rho_1 x_{h,N})$ is always larger than a strictly positive number independent of N. Therefore, if (1.A.4) does not hold, then

$$KL(\frac{\hat{T}_N}{N}, \rho_1 x_{h,N}) - KL(\frac{\hat{T}_N}{N}, \rho_2 x_{h,N})$$

is always smaller than a positive number independent of N. Hence, if we cannot find an $\hat{N}_1$ such that (1.A.4) holds when $N > \hat{N}_1$, then from (1.A.2), for each $\hat{M}_1 > 0$, we can find $\bar{N}_1$ such that

$$\frac{\mathbb{P}[\hat{T}_{\bar{N}_1}; \bar{N}_1 | \theta_2]}{\mathbb{P}[\hat{T}_{\bar{N}_1}; \bar{N}_1 | \theta_1]} > \hat{M}_1,$$

and

$$\frac{\mathbb{P}[\hat{T}_{\bar{N}_1} - 1; \bar{N}_1 | \theta_2]}{\mathbb{P}[\hat{T}_{\bar{N}_1} - 1; \bar{N}_1 | \theta_1]} > \hat{M}_1,$$

By choosing $\hat{M}_1$ large enough, we can see that the principal strictly prefers A when he observes $\hat{T}_{\bar{N}_1} - 1$ approvals from $\bar{N}_1$ agents, which contradicts the optimality of $\hat{T}_{\bar{N}_1}$ as shown in (1.8).

However, if there exists $\hat{N}_1$ such that (1.A.4) holds when $N > \hat{N}_1$, we must have

$$\lim_{N\to\infty} \frac{\mathbb{P}[\hat{T}_N; N | \theta_2]}{\mathbb{P}[\hat{T}_N; N | \theta_3]} = \infty. \tag{1.A.5}$$

Note that when $\hat{T}_N < N \cdot \rho_2 x_{h,N}$,

$$\begin{aligned} KL(\frac{\hat{T}_N}{N}, \rho_3 x_{h,N}) - KL(\frac{\hat{T}_N}{N}, \rho_2 x_{h,N}) &= \frac{\hat{T}_N}{N} \log \frac{\rho_2 x_{h,N}}{\rho_3 x_{h,N}} + (1 - \frac{\hat{T}_N}{N}) \log \frac{1-\rho_2 x_{h,N}}{1-\rho_3 x_{h,N}} \\ &\geq \rho_2 x_{h,N} \cdot \log \frac{\rho_2 x_{h,N}}{\rho_3 x_{h,N}} + (1-\rho_2 x_{h,N}) \log \frac{1-\rho_2 x_{h,N}}{1-\rho_3 x_{h,N}} \\ &= KL(\rho_2 x_{h,N}, \rho_3 x_{h,N}). \\ &> 0. \end{aligned}$$

From (1.A.3), we can see that $KL(\rho_2 x_{h,N}, \rho_3 x_{h,N})$ is always larger than a strictly positive number independent of N. We then prove (1.A.5) by (1.A.2).

From (1.A.5),

$$\lim_{N\to\infty} \frac{\mathbb{P}[\hat{T}_N - 1; N|\theta_3]}{\mathbb{P}[\hat{T}_N - 1; N|\theta_2]} = 0,$$

Hence, the agents choose $x_{h,N} = x_{l,N} = 0$ when N is above some threshold by (1.6), which leads to a contradiction.

Therefore, in every sequence of informative equilibrium $\{\Gamma_N = (x_{h,N}, \hat{T}_N)\}$, we must have

$$\lim_{N\to\infty} x_{h,N} = 0. \tag{1.A.6}$$

Otherwise, we can construct a subsequence from it and show that x_h converges to a positive number along this sub-sequence, which leads to a contraction as shown above.

We now assume that there exists a sequence of informative equilibrium $\{\Gamma_N = (x_{h,N}, \hat{T}_N)\}$ with

$$\lim_{N\to\infty} N \cdot x_{h,N} = \infty. \tag{1.A.7}$$

From the proof above, we have

$$KL(\frac{\hat{T}_N}{N}, \rho_1 x_{h,N}) - KL(\frac{\hat{T}_N}{N}, \rho_2 x_{h,N}) \geq KL(\rho_2 x_{h,N}, \rho_1 x_{h,N}), \text{ if } \hat{T}_N \geq N \cdot \rho_2 x_{h,N}, \tag{1.A.8}$$

$$KL(\frac{\hat{T}_N}{N}, \rho_3 x_{h,N}) - KL(\frac{\hat{T}_N}{N}, \rho_2 x_{h,N}) \geq KL(\rho_2 x_{h,N}, \rho_3 x_{h,N}),, \text{ if } \hat{T}_N \leq N \cdot \rho_2 x_{h,N}. \tag{1.A.9}$$

We can linearize $KL(\rho_2 x_{h,N}, \rho_i x_{h,N}))$ with respect to $x_{h,N}$ when $x_{h,N} \approx 0$ for $i \in \{1,3\}$,

$$KL(\rho_2 x_{h,N}, \rho_i x_{h,N}) = \kappa_i x_{h,N} + o(x_{h,N}), \ \forall i \in \{1,3\}, \tag{1.A.10}$$

where

$$\kappa_i = \rho_2 \log \frac{\rho_2}{\rho_i} + \rho_i - \rho_2 > 0, \ \forall i \in \{1,3\}.$$

Similar to the proof before, when (1.A.6) and (1.A.7) hold, from (1.A.2), (1.A.8) and (1.A.10), when N is above some threshold, we must have (1.A.4), which leads to (1.A.5) from (1.A.2), (1.A.9) and (1.A.10), and a further contradiction.

Therefore, there exists a finite number T_0 independent to N such that for each N and each informative equilibrium with N agnets,

$$N \cdot x_h < T_0.$$

Otherwise, we can construct a sequence of informative equilibria and construct a subsequence from it, showing that $N \cdot x_{h,N}$ grows without bound along this subsequence, which leads to a contradiction as shown above.

1.A.1.3 Proof of Proposition 1.3

Consider a sequence of informative equilibrium $\{\Gamma_N = (x_{h,N}, \hat{T}_N)\}$. From Proposition 1.2,

$$x_{h,N} < \frac{T_0}{N}, \ \forall N \in \mathbb{N}^+.$$

Hence,

$$\frac{\mathbb{P}[T;N|\theta_3]}{\mathbb{P}[T;N|\theta_1]} = \frac{(\rho_3 x_{h,N})^T(1-\rho_3 x_{h,N})^{N-T}}{(\rho_1 x_{h,N})^T(1-\rho_1 x_{h,N})^{N-T}} > \left(\frac{\rho_3}{\rho_1}\right)^T \left(\frac{1-\rho_3\cdot\frac{T_0}{N}}{1-\rho_1\cdot\frac{T_0}{N}}\right)^{N-T} > \left(\frac{\rho_3}{\rho_1}\right)^T \left(\frac{1-\rho_3\cdot\frac{T_0}{N}}{1-\rho_1\cdot\frac{T_0}{N}}\right)^{N}.$$

Since

$$\lim_{N\to\infty}\left(\frac{1-\rho_3\cdot\frac{T_0}{N}}{1-\rho_1\cdot\frac{T_0}{N}}\right)^{N} = \exp[(\rho_1-\rho_3)T_0] > 0.$$

We can find $\gamma > 0$ independent of T and N such that

$$\frac{\mathbb{P}[T;N|\theta_3]}{\mathbb{P}[T;N|\theta_1]} > \left(\frac{\rho_3}{\rho_1}\right)^T \cdot \gamma, \ \forall N \in \mathbb{N}^+ \text{ and } \forall T \in \{0, \ldots, N\}. \tag{1.A.11}$$

Note that $\frac{\mathbb{P}[\hat{T}_N-1;N|\theta_3]}{\mathbb{P}[\hat{T}_N-1;N|\theta_1]}$ must be always smaller than a number independent of N for each $N \in \mathbb{N}^+$. Otherwise, the principal chooses A when he observes $\hat{T}_N - 1$ approvals, which contradicts the optimality of $\hat{T}_N$ as shown in (1.8). From (1.A.11), for each $N \in \mathbb{N}^+$, the equilirium cut-off $\hat{T}_N$ is always smaller than a number indepenedent of N.

1.A.2 Informative Equilibrium with $\hat{T} = 1$

We provide sufficient and necessary conditions under which there exists an informative equilibrium with $\hat{T} = 1$, that is, the equilibrium in which the principal chooses the unanimity rule. We then extend this approach for the informative equilibrium with $\hat{T} = i$ for each $i \in \mathbb{N}$.

We only consider the case where

$$\frac{V_{ag}(\theta_1)}{V_{ag}(\theta_3)} \cdot \frac{V_{pc}(\theta_3)}{V_{pc}(\theta_1)} < \frac{\rho_3}{\rho_1}.$$

Otherwise, from Proposition 1.4, information transmission fails when N is large enough.

For each $i \in \mathbb{N}^+$, we say ***information transmission persists with $\hat{T} = i$*** (ITP_i) if there exists $\hat{N}_1$ such that for each $N > \hat{N}_1$, an informative equilibrium with $\hat{T} = i$ exists. We say ***information transmission fails with $\hat{T} = i$*** (ITF_i) if there exists $\hat{N}_2$ such that for each $N > \hat{N}_2$, there does not exist an informative equilibrium with $\hat{T} = i$.

1.A.2.1 Unanimity Rule

We now provide the sufficient and necessary conditions for ITP_1 and IFP_1. We consider the case where the principal always chooses $\hat{T} = 1$, that is, he follows the unanimity rule under which B is chosen if all the agents reject it. From Proposition 1.2, when N is large enough, the agents must choose $x_h < 1$ in each informative equilibrium. Therefore, if ITP_1 holds, then there exists $\hat{N}_1$ and a sequence of informative equilibrium $\{\Gamma_N = (x_{h,N}, x_{\ell,N}, \hat{T})\}_{N=\hat{N}_1}^{\infty}$ with

$$x_{h,N} \in (0,1), x_{\ell,N} = 0, \hat{T} = 1,$$

for each $N > \hat{N}_1$. We suppress $x_{h,N}$ to x_N to save notation. If ITP_1 holds, there exists a strictly positive sequence $\{x_N\}_{N=\hat{N}_1}^{\infty}$ satisfying

$$\frac{\rho_3 \cdot (1-\rho_3 x_N)^{N-1} \cdot V_{ag}(\theta_3)}{-\lambda_2 \cdot \rho_2 \cdot (1-\rho_2 x_N)^{N-1} \cdot V_{ag}(\theta_2) - \lambda_1 \cdot \rho_1 \cdot (1-\rho_1 x_N)^{N-1} \cdot V_{ag}(\theta_1)} = 1, \tag{1.A.12}$$

$$\frac{\rho_3 \cdot (1-\rho_3 x_N)^{N-1} \cdot V_{pc}(\theta_3) + \lambda_2 \cdot \rho_2 \cdot (1-\rho_2 x_N)^{N-1} \cdot V_{pc}(\theta_2)}{-\lambda_1 \cdot \rho_1 \cdot (1-\rho_1 x_N)^{N-1} \cdot V_{pc}(\theta_1)} > 1, \tag{1.A.13}$$

$$\frac{(1-\rho_3 x_N)^{N} \cdot V_{pc}(\theta_3) + \lambda_2 \cdot (1-\rho_2 x_N)^{N} \cdot V_{pc}(\theta_2)}{-\lambda_1 \cdot (1-\rho_1 x_N)^{N} \cdot V_{pc}(\theta_1)} \leq 1. \tag{1.A.14}$$

for each $N > \hat{N}_1$. We derive (1.A.12) from (1.6) since the agents are indifferent conditional on receiving signal h and being pivotal. We derive (1.A.13) and (1.A.14) from (1.6) since the principal prefers A when one agent approve it and prefers B when all the agents reject it.

Note that we can interpret the agents as voters who vote between A and B under the unanimity rule. An informative voting equilibrium under the unanimity rule exists if there is a positive x_N satisfying (1.A.12). We now provide conditions under which there exists $\hat{N}_1$ such that when $N > \hat{N}_1$, there exists an informative equilibrium under the unanimity rule. Note that the left side of (1.A.12) decreases with x_N. Furthermore, for each $x \in (0,1)$, if we fix $x_N = x$ for each N and let N go to infinity, the left side approaches 0. Therefore, by the intermediate value theorem, a sufficient and necessary condition under which informative voting equilibria exist when N is large enough is that the value of the left side (1.A.12) given that $x_N = 0$ is strictly bigger than 1, by which we have

$$\lambda_2 < -\frac{\rho_3 V_{ag}(\theta_3) + \lambda_1 \rho_1 V_{ag}(\theta_1)}{\rho_2 V_{ag}(\theta_2)}.$$

Denote the right side by $\hat{\lambda}'_{2,1}$. Note that if $\lambda_2 < \hat{\lambda}'_{2,1}$. There exists a unique informative voting equilibrium since (1.A.12) admits a unique solution. Therefore, a necessary condition for ITP_1 is $\lambda_2 < \hat{\lambda}'_{2,1}$ while a sufficient condition for ITF_1 is $\lambda_2 \geq \hat{\lambda}'_{2,1}$.

Now consider the case where $\lambda_2 < \hat{\lambda}'_{2,1}$, there exists $\hat{N}_1$ such that there exists a strictly positive sequence $\{x_N\}_{N=\hat{N}_1}^{\infty}$ satisfying (1.A.12) for each $N > \hat{N}_1$. Note that (1.A.12) implies (1.A.13) since

$$V_{pc}(\theta_2) > 0 > V_{pc}(\theta_2),$$

$$-\frac{V_{pc}(\theta_1)}{V_{pc}(\theta_3)} \leq -\frac{V_{ag}(\theta_1)}{V_{ag}(\theta_3)}.$$

Assume that there exists $\hat{N}_1$ such that there exists a strictly positive sequence $\{x_N\}_{N=\hat{N}_1}^{\infty}$ satisfying (1.A.12) and (1.A.14) for each $N > \hat{N}_1$, by which we have ITP_1. By Proposition 1.2,

$$\lim_{N\to\infty} x_N = 0. \tag{1.A.15}$$

Therefore, we can find $\{\epsilon_N\}_{\hat{N}_1}^{\infty}$ with $\lim_{N\to\infty} \epsilon_N = 0$ such that for each $N > \hat{N}_1$

$$\frac{(1-\rho_3 x_N)^{N-1} \cdot V_{pc}(\theta_3) + \lambda_2 \cdot (1-\rho_2 x_N)^{N-1} \cdot V_{pc}(\theta_2)}{-\lambda_1 \cdot (1-\rho_1 x_N)^{N-1} \cdot V_{pc}(\theta_1)} < \frac{1}{1-\epsilon_N} \tag{1.A.16}$$

by (1.A.14) and (1.A.15).

Note that we can rewrite (1.A.12) as

$$a_1 \left(\frac{1-\rho_1 x_N}{1-\rho_3 x_N}\right)^{N-1} + a_2 \left(\frac{1-\rho_2 x_N}{1-\rho_3 x_N}\right)^{N-1} = 1, \tag{1.A.17}$$

where $a_1 > 0$ and $a_2 > 0$ are calculated from (1.A.12). We can also rewrite (1.A.16) as

$$b_1 \left(\frac{1-\rho_1 x_N}{1-\rho_3 x_N}\right)^{N-1} - b_2 \left(\frac{1-\rho_2 x_N}{1-\rho_3 x_N}\right)^{N-1} > 1-\epsilon_N, \tag{1.A.18}$$

where $b_1 > 0$ and $b_2 > 0$ are calculated from (1.A.16). By (1.A.17) and (1.A.18),

$$\left(\frac{1-\rho_1 x_N}{1-\rho_3 x_N}\right)^{N-1} > \frac{b_2 + a_2(1-\epsilon_N)}{a_2 b_1 + a_1 b_2}, \tag{1.A.19}$$

$$\left(\frac{1-\rho_2 x_N}{1-\rho_3 x_N}\right)^{N-1} < \frac{b_1 - a_1(1-\epsilon_N)}{a_2 b_1 + a_1 b_2}. \tag{1.A.20}$$

Note that for each $t > 0$, if there exists a sequence $\{y_N\}_{N=\hat{N}_1}^{\infty}$ such that

$$\left(\frac{1-\rho_1 y_N}{1-\rho_3 y_N}\right)^{N-1} > t, \ \forall N > \hat{N}_1.$$

then

$$\liminf_{N\to\infty}\left(\frac{1-\rho_2 y_N}{1-\rho_3 y_N}\right)^{N-1} > t^{\frac{\rho_3-\rho_2}{\rho_3-\rho_1}},$$

which is shown by considering the sequence $\{y_N\}_{N=\hat{N}_1}^{\infty}$ such that

$$\left(\frac{1-\rho_1 y_N}{1-\rho_3 y_N}\right)^{N-1} = t, \ \forall N > \hat{N}_1.$$

Therefore, if both (1.A.19) and (1.A.20) for each x_N when $N > \hat{N}_1$ with $\lim_{N\to\infty} \epsilon_N = 0$, we must have

$$\left(\frac{b_2 + a_2}{a_2 b_1 + a_1 b_2}\right)^{\frac{\rho_3-\rho_2}{\rho_3-\rho_1}} \leq \frac{b_1 - a_1}{a_2 b_1 + a_1 b_2}$$

That is,

$$\lambda_2 \leq \lambda_1^{\frac{\rho_3-\rho_2}{\rho_3-\rho_1}} \cdot \frac{\rho_3 u_{13} - \rho_1 v_{13}}{(\rho_3 u_{23} + \rho_2 v_{23})^{\frac{\rho_3-\rho_2}{\rho_3-\rho_1}} (\rho_1 u_{23} v_{13} + \rho_2 u_{13} v_{23})^{\frac{\rho_2-\rho_1}{\rho_3-\rho_1}}}.$$

with

$$u_{ij} = \frac{|V_{pc}(\theta_i)|}{|V_{pc}(\theta_j)|}, \ \forall i,j \in \{1,2,3\},$$

$$v_{ij} = \frac{|V_{ag}(\theta_i)|}{|V_{ag}(\theta_j)|}, \ \forall i,j \in \{1,2,3\}.$$

Denote the right side by $\hat{\lambda}''_{2,1}$. A nessessary condition for ITP_1 is $\lambda_2 \leq \hat{\lambda}''_{2,1}$. Furthermore, if $\lambda_2 > \hat{\lambda}''_{2,1}$, we can find $\hat{N}_2$ such that for each $N > \hat{N}_2$, there does not exist x_N satisfying both (1.A.19) and (1.A.20). Therefore, a sufficiecnt condition for ITF_1 is $\lambda_2 > \hat{\lambda}''_{2,1}$.

Let

$$\hat{\lambda}_{2,1} = \min\{\hat{\lambda}'_{2,1}, \hat{\lambda}''_{2,1}\}.$$

A nessessary condition for ITP_1 is $\lambda_2 \leq \hat{\lambda}_{2,1}$ while a sufficient condition for ITF_1 is $\lambda_2 > \hat{\lambda}_{2,1}$.

We now show that a sufficient condition for ITP_1 is $\lambda_2 < \hat{\lambda}_{2,1}$. Note that we can find sequences $\{y_N\}_{N=1}^{\infty}$, $\{\epsilon_N\}_{N=1}^{\infty}$, and $\{\epsilon'_N\}_{N=1}^{\infty}$ with

$$\lim_{N\to\infty} y_N = \lim_{N\to\infty} \epsilon_N = \lim_{N\to\infty} \epsilon'_N = 0,$$

such that

$$\frac{\rho_3 \cdot (1-\rho_3 y_N)^{N-1} \cdot V_{ag}(\theta_3)}{-\hat{\lambda}''_{2,1} \cdot \rho_2 \cdot (1-\rho_2 y_N)^{N-1} \cdot V_{ag}(\theta_2) - \lambda_1 \cdot \rho_1 \cdot (1-\rho_1 y_N)^{N-1} \cdot V_{ag}(\theta_1)} = 1+\epsilon_N,$$

$$\frac{(1-\rho_3 y_N)^N \cdot V_{pc}(\theta_3) + \hat{\lambda}''_{2,1} \cdot (1-\rho_2 y_N)^N \cdot V_{pc}(\theta_2)}{-\lambda_1 \cdot (1-\rho_1 y_N)^N \cdot V_{pc}(\theta_1)} = 1+\epsilon'_N$$

Hence, for each $\lambda_2 < \hat{\lambda}_{2,1}$, we can find $\hat{N}_1$ such that for each $N > \hat{N}_1$,

$$\frac{\rho_3 \cdot (1-\rho_3 y_N)^{N-1} \cdot V_{ag}(\theta_3)}{\lambda_2 \cdot \rho_2 \cdot (1-\rho_2 y_N)^{N-1} \cdot V_{ag}(\theta_2) - \lambda_1 \cdot \rho_1 \cdot (1-\rho_1 y_N)^{N-1} \cdot V_{ag}(\theta_1)} > 1,$$

$$\frac{(1-\rho_3 y_N)^N \cdot V_{pc}(\theta_3) + \lambda_2 \cdot (1-\rho_2 y_N)^N \cdot V_{pc}(\theta_2)}{-\lambda_1 \cdot (1-\rho_1 y_N)^N \cdot V_{pc}(\theta_1)} < 1.$$

Note that the left sides of both equations above decrease with y_N. Therefore, for each $N > \hat{N}_1$, if we can find $x_N \le 1$ satisfying (1.A.12), then we have $x_N > y_N$ and (1.A.14) is satisfied. In this case, since $\lambda_2 < \hat{\lambda}_{2,1} \le \hat{\lambda}_{2,1}$ and x_N is the unique solution of (1.A.12), we must have $x_N > 0$. Therefore, we construct an informative equilibrium with $\hat{T} = 1$. If we cannot find $x_N \le 1$ satisfying (1.A.12), we have

$$\frac{\rho_3 \cdot (1-\rho_3)^{N-1} \cdot V_{ag}(\theta_3)}{\lambda_2 \cdot \rho_2 \cdot (1-\rho_2)^{N-1} \cdot V_{ag}(\theta_2) - \lambda_1 \cdot \rho_1 \cdot (1-\rho_1)^{N-1} \cdot V_{ag}(\theta_1)} > 1, \quad (1.A.21)$$

$$\frac{(1-\rho_3)^N \cdot V_{pc}(\theta_3) + \lambda_2 \cdot (1-\rho_2)^N \cdot V_{pc}(\theta_2)}{-\lambda_1 \cdot (1-\rho_1)^N \cdot V_{pc}(\theta_1)} < 1. \quad (1.A.22)$$

From (1.A.21),

$$\frac{\rho_3 \cdot (1-\rho_3)^{N-1} \cdot V_{pc}(\theta_3) + \lambda_2 \cdot \rho_2 \cdot (1-\rho_2)^{N-1} \cdot V_{pc}(\theta_2)}{-\lambda_1 \cdot \rho_1 \cdot (1-\rho_1)^{N-1} \cdot V_{pc}(\theta_1)} < 1. \quad (1.A.23)$$

From (1.A.21),

$$\frac{(1-\rho_3) \cdot (1-\rho_3)^{N-1} \cdot V_{ag}(\theta_3)}{\lambda_2 \cdot (1-\rho_2) \cdot (1-\rho_2)^{N-1} \cdot V_{ag}(\theta_2) - \lambda_1 \cdot (1-\rho_1) \cdot (1-\rho_1)^{N-1} \cdot V_{ag}(\theta_1)} < 1, \quad (1.A.24)$$

From (1.A.21) and (1.A.24), the agents strictly prefer A conditional receiving signal h and being pivotal while strictly prefer B conditional on receiving signal ℓ and being pivotal. Hence, it is optimal for the agents to choose $x_h = 1$ and $x_\ell = 0$. From (1.A.22) and (1.A.23), it is optimal for the principal to choose $\hat{T} = 1$. Hence, we construct an informative equilibrium. Therefore, for each $\lambda_2 < \hat{\lambda}_{2,1}$, we can find $\hat{N}_1$ such that for each $N > \hat{N}_1$, an informative equilibrium with $\hat{T} = 1$ exists.

1.A.2.2 General Case

We now discuss the sufficient conditions for ITP_i and IFP_i for each $i > 1$. Fix an arbitrary $i > 1$ and consider the case where the principal always chooses $\hat{T} = i$. From Proposition 1.2, when N is large enough, the agents must choose $x_h < 1$ in each informative equilibrium. Therefore, if ITP_i holds, then there exists $\hat{N}_1$ and a sequence of informative equilibrium $\{\Gamma_N = (x_{h,N}, x_{\ell,N}, \hat{T})\}_{N=\hat{N}_1}^{\infty}$ with

$$x_{h,N} \in (0,1), x_{\ell,N} = 0, \hat{T} = i,$$

for each $N > \hat{N}_1$. We suppress $x_{h,N}$ to x_N to save notation. If ITP_i holds, there exists a strictly positive sequence $\{x_N\}_{N=\hat{N}_1}^{\infty}$ satisfying

$$\frac{\rho_3 \cdot (\rho_3 x_N)^{i-1} \cdot (1-\rho_3 x_N)^{N-i} \cdot V_{ag}(\theta_3)}{-\lambda_2 \cdot \rho_2 \cdot (\rho_2 x_N)^{i-1} \cdot (1-\rho_2 x_N)^{N-i} \cdot V_{ag}(\theta_2) - \lambda_1 \cdot \rho_1 \cdot (\rho_1 x_N)^{i-1} \cdot (1-\rho_1 x_N)^{N-i} \cdot V_{ag}(\theta_1)} = 1, \tag{1.A.25}$$

$$\frac{\rho_3 \cdot (\rho_3 x_N)^{i} \cdot (1-\rho_3 x_N)^{N-i} \cdot V_{pc}(\theta_3) + \lambda_2 \cdot \rho_2 \cdot (\rho_2 x_N)^{i} \cdot (1-\rho_2 x_N)^{N-i} \cdot V_{pc}(\theta_2)}{-\lambda_1 \cdot \rho_1 \cdot (\rho_1 x_N)^{i} \cdot (1-\rho_1 x_N)^{N-i} \cdot V_{pc}(\theta_1)} > 1, \tag{1.A.26}$$

$$\frac{(\rho_3 x_N)^{i-1} \cdot (1-\rho_3 x_N)^{N-i+1} \cdot V_{pc}(\theta_3) + \lambda_2 \cdot (\rho_3 x_N)^{i-1} \cdot (1-\rho_2 x_N)^{N-i+1} \cdot V_{pc}(\theta_2)}{-\lambda_1 \cdot (\rho_3 x_N)^{i-1} \cdot (1-\rho_1 x_N)^{N-i+1} \cdot V_{pc}(\theta_1)} \leq 1. \tag{1.A.27}$$

Note that we can choose

$$\lambda_1' = \lambda_2 \cdot \left(\frac{\rho_1}{\rho_3}\right)^{i-1},$$

$$\lambda_2' = \lambda_2 \cdot \left(\frac{\rho_2}{\rho_3}\right)^{i-1},$$

and convert (1.A.25), (1.A.26), and (1.A.27) regading λ_1 and λ_2 to (1.A.12), (1.A.13), and (1.A.14) regarding λ_1' and λ_2'. In this way, we can follow the analysis in A.2.1 and find $\hat{\lambda}_{2,i}$ such that ITP_i holds if $\lambda_2 < \hat{\lambda}_{2,i}$ and while ITF_i holds if $\lambda_2 > \hat{\lambda}_{2,i}$. The left panel of Figure 1.8 illustrateS $\hat{\lambda}_{2,1}$, $\hat{\lambda}_{2,2}$, and $\hat{\lambda}_{2,3}$.

1.A.3 Proof for the Results in Section 1.4

1.A.3.1 Proof of Proposition 4

Assume that there exists an informative equilibrium with $x_h \in (0,1)$ and $x_\ell = 0$, that is, the agents are indifferent conditional on receiving signal h and being pivotal. From (1.6),

$$\frac{q_3^0 \cdot \rho_3 \cdot \mathbb{P}[T-1;N-1|\theta_3] \cdot V_{ag}(\theta_3)}{-\sum_{i=1}^{2} q_i^0 \cdot \rho_i \cdot \mathbb{P}[T-1;N-1|\theta_i] \cdot V_{ag}(\theta_i)} = 1.$$

Therefore,

$$\frac{q_3^0 \cdot \rho_3 \cdot \mathbb{P}[T-1;N-1|\theta_3] \cdot V_{ag}(\theta_3)}{-q_1^0 \cdot \rho_1 \cdot \mathbb{P}[T-1;N-1|\theta_1] \cdot V_{ag}(\theta_1)} > 1,$$

$$\frac{q_3^0 \cdot \mathbb{P}[T-1;N-1|\theta_3]}{q_1^0 \cdot \mathbb{P}[T-1;N-1|\theta_1]} > -\frac{\rho_1}{\rho_3}\frac{V_{ag}(\theta_1)}{V_{ag}(\theta_3)}. \tag{1.A.28}$$

Furthermore, since the principal must prefer B when he observes $\hat{T}-1$ approvals by (1.8),

$$\frac{\sum_{i=2}^{3} q_i^0 \cdot \mathbb{P}[T-1;N|\theta_i] \cdot V_{pc}(\theta_i)}{-q_1^0 \cdot \mathbb{P}[T-1;N|\theta_1] \cdot V_{pc}(\theta_1)} \leq 1.$$

Therefore,

$$\frac{q_3^0 \cdot \mathbb{P}[T-1;N|\theta_3] \cdot V_{pc}(\theta_3)}{-q_1^0 \cdot \mathbb{P}[T-1;N|\theta_1] \cdot V_{pc}(\theta_1)} \leq 1,$$

$$\frac{q_3^0 \cdot (1-\rho_3 x_h) \cdot \mathbb{P}[T-1;N-1|\theta_3] \cdot V_{pc}(\theta_3)}{-q_1^0 \cdot (1-\rho_1 x_h) \cdot \mathbb{P}[T-1;N|\theta_1] \cdot V_{pc}(\theta_1)} \leq 1,$$

$$\frac{\sum_{i=2}^{3} q_i^0 \cdot \mathbb{P}[T-1;N|\theta_i] \cdot V_{pc}(\theta_i)}{-q_1^0 \cdot \mathbb{P}[T-1;N-1|\theta_1] \cdot V_{pc}(\theta_1)} \leq -\frac{1-\rho_1 x_h}{1-\rho_3 x_h}\frac{V_{pc}(\theta_1)}{V_{pc}(\theta_3)}. \tag{1.A.29}$$

From (1.A.28) and (1.A.29),

$$-\frac{\rho_1}{\rho_3}\frac{V_{ag}(\theta_1)}{V_{ag}(\theta_3)} < -\frac{1-\rho_1 x_h}{1-\rho_3 x_h}\frac{V_{pc}(\theta_1)}{V_{pc}(\theta_3)},$$

$$\frac{V_{ag}(\theta_1)}{V_{ag}(\theta_3)} \cdot \frac{V_{pc}(\theta_3)}{V_{pc}(\theta_1)} < \frac{\rho_3}{\rho_1} \cdot \frac{1-\rho_1 x_h}{1-\rho_3 x_h}. \tag{1.A.30}$$

By Proposition 1.2, when N is large enough, in every informative equilibrium, we must have $x_h \approx 0$. Therefore, if

$$\frac{V_{ag}(\theta_1)}{V_{ag}(\theta_3)} \cdot \frac{V_{pc}(\theta_3)}{V_{pc}(\theta_1)} > \frac{\rho_3}{\rho_1},$$

there does not exist any informative equilibrium when N is large enough, by which we prove Proposition 1.4.

1.A.3.2 Proof of Proposition 1.5

We choose

$$\hat{\lambda}_2 = \sup_{i \in \mathbb{N}+} \hat{\lambda}_{2,i}.$$

From the analysis in Appendix 1.A.2, it is direct to see that information transmission persists if $\lambda_2 < \hat{\lambda}_2$.

For each $\lambda_2 > \hat{\lambda}_2$, by Proposition 1.3, we can find T_0 independent of N such that in any informative equilibrium, the principal chooses $\hat{T} < T_0$. We further have

$$\lambda_2 > \max_{i \in \mathbb{N}+} \hat{\lambda}_{2,i} > \max_{i < T_0} \hat{\lambda}_{2,i}.$$

Hence, we have ITF_i for each $i < T_0$. We then can find $\hat{N}_2$ such that when $N > \hat{N}_2$, only the babbling equilibrium exists since T_0 is a finite number independent to N, by which information transmission fails.

1.A.3.3 Rest Results of Section 1.4

The proofs of Corollary 1.2, Corollary 1.3, and Proposition 1.7 are similar. We only need to show that given that there exists an informative equilibrium, there always exists an informative equilibrium with a higher x_h when there are a lower λ_2, a lower ρ_1, higher $V_{ag}(\theta_i)$, and lower $V_{pc}(\theta_i)$ for each $i \in \{1, 2, 3\}$. We skip the proof here since we already do a similar construction when proving a sufficient condition for ITP_1 is $\lambda_2 < \hat{\lambda}''_{2,1}$ in Appendix 1.A.2.

The part of Corollary 1.3 that $\hat{\lambda}_2$ is non-monotonic with ρ_2 and ρ_3 is proved by taking the derivative of $\hat{\lambda}_2$ over ρ_2 and ρ_3. Proposition 1.6 is direct from Blackwell's informative ranking. Proposition 1.8 is based on calculation, which can be directly seen by the fact that $\hat{\lambda}''_{2,1}$ is concave in λ_1.

1.A.4 Mixed-Strategy Equilibria

In this section, we characterize the informative equilibrium in which the principal uses a mixed strategy. Consider an equilibrium in which the agents choose an informative strategy $\boldsymbol{x}$ with $x_\ell < x_h$ and the principal chooses $\hat{T} \in \{1, \ldots, N-1\}$ and $p \in (0, 1)$ such that he chooses B when $T < \hat{T}$, chooses A with probability p when $T = \hat{T}$, and chooses A when $T > \hat{T}$.[33]

33. We skip the case where $\hat{T} = 0$ and $p \in (0, 1)$ since if there exists such an equilibrium, there must exist one informative with $\hat{T} = 1$ and $p = 1$. Similarly, we skip the case where $\hat{T} = N$ and $p \in (0, 1)$.

In this case, if one agent is pivotal, with probability p, there are $\hat{T}-1$ approvals from $N-1$ agents while with probability $1-p$, there are $\hat{T}$ approvals from $N-1$ agents. Hence,

$$\mathbb{P}[piv|\theta_i;\boldsymbol{x},\hat{T},p] = p \cdot \mathbb{P}[\hat{T}-1,N-1|\theta_i;\boldsymbol{x},\hat{T},p] + (1-p)\cdot \mathbb{P}[\hat{T},N-1|\theta_i;\boldsymbol{x},\hat{T},p].$$

We then define

$$L_{ag}(s;\boldsymbol{x},\hat{T},p) = \frac{q_3^0 \cdot \mathbb{P}[s|\theta_3]\cdot \mathbb{P}[piv|\theta_3;\boldsymbol{x},\hat{T},p]\cdot V_{ag}(\theta_3)}{-\sum_{i=1}^{2} q_i^0 \cdot \mathbb{P}[s|\theta_i]\cdot \mathbb{P}[piv|\theta_i;\boldsymbol{x},\hat{T},p]\cdot V_{ag}(\theta_i)}.$$

We have

$$\begin{cases} x_s = 1 & \text{when } L_{ag}(s;\boldsymbol{x},\hat{T},p) > 1, \\ x_s \in [0,1] & \text{when } L_{ag}(s;\boldsymbol{x},\hat{T},p) = 1, \\ x_s = 0 & \text{when } L_{ag}(s;\boldsymbol{x},\hat{T},p) < 1. \end{cases} \tag{1.A.31}$$

For the principal, he must be indifferent when he observes $\hat{T}$,

$$\frac{\sum_{i=2}^{3} q_i^0 \cdot \mathbb{P}[\hat{T};N|\theta_i]\cdot V_{pc}(\theta_i)}{-q_1^0 \cdot \mathbb{P}[\hat{T};N|\theta_1]\cdot V_{ag}(\theta_1)} = 1. \tag{1.A.32}$$

Therefore, an informative equilibrium $\{\boldsymbol{x},\hat{T},p\}$ with $p\in(0,1)$ is characterized by (1.A.31) and (1.A.32).

It is direct to verify that Lemma 1.1, Proposition 1.2, Proposition 1.3, and hence Theorem 1.1 stay valid when we allow the principal to use a mixed strategy. We extend results in Section 1.4 based on the following lemma.

Lemma 1.3. *If there exists an informative equilibrium in which the principal chooses A with probability $p\in(0,1)$ when he observes $\hat{T}$ approvals with $\hat{T}\in\{1,\ldots,N-1\}$, then there exists an informative equilibrium in which the principal chooses A if and only if $T>\hat{T}$.*

Proof. Consider the case that there exists an informative equilibrium $\{x_h,\hat{T},p\}$ in which the principal chooses A with probability $p\in(0,1)$ when he observes $\hat{T}$ approvals. Since the agents chooses $x_h>0$, from (1.A.31), we have

$$\frac{q_3^0 \cdot \rho_3 \cdot \mathbb{P}[piv|\theta_3;\boldsymbol{x},\hat{T},p]\cdot V_{ag}(\theta_3)}{-\sum_{i=1}^{2} q_i^0 \cdot \rho_i \cdot \mathbb{P}[piv|\theta_i;\boldsymbol{x},\hat{T},p]\cdot V_{ag}(\theta_i)} \leq 1. \tag{1.A.33}$$

Note that since $x_\ell < x_h$, we have

$$\frac{\mathbb{P}[T;N|\theta_3]}{\mathbb{P}[T;N|\theta_1]} \text{ and } \frac{\mathbb{P}[T;N|\theta_2]}{\mathbb{P}[T;N|\theta_1]}$$

both strictly increase with T. Hence, from (1.A.33),

$$\frac{q_3^0 \cdot \rho_3 \cdot \mathbb{P}[\hat{T}, N-1|\theta_3; \boldsymbol{x}, \hat{T}, p] \cdot V_{ag}(\theta_3)}{-\sum_{i=1}^{2} q_i^0 \cdot \rho_i \cdot \mathbb{P}[\hat{T}, N-1|\theta_i; \boldsymbol{x}, \hat{T}, p] \cdot V_{ag}(\theta_i)} > 1. \tag{1.A.34}$$

Note that if $x_h = 1$, equations (1.A.32) and (1.A.34) guarantee that there exists an information equilibrium in which the agents choose the same x_h and the principal chooses A if and only if $T > \hat{T}$. Note that (1.A.32) implies that the agents must choose $x_\ell = 0$ conditional on receiving signal ℓ and being pivotal. If $x_h < 1$, we can follow the construction when proving a sufficient condition for ITP_1 is $\lambda_2 < \hat{\lambda}''_{2,1}$ in Appendix 1.A.2, showing that there must exist $x'_h \in (x_h, 1]$ such that there exists an informative equilibrium in which the agents choose the same x'_h and the principal chooses A if and only if $T > \hat{T}$.

□

From Lemma 1.3, we can see that we only need to focus on the informative equilibrium in which the principal follows pure strategy when discussing the existence of informative equilibria.

1.A.5 Commitment Case

1.A.5.1 Proof of Proposition 1.10

Fix $N \in \mathbb{N}^+$ and conider a mechanism $M_N = (\mu_N, \hat{T}_N^\alpha, \hat{T}_N^\beta,)$. When one agent is pivotal, with probability μ_N, there are $\hat{T}_N^\alpha$ agents receving signal h for $N-1$ agents, and with probability $1-\mu_N$, there are $\hat{T}_N^\beta$ agents receving signal h for $N-1$ agents.

Define

$$\mathbb{P}[piv_\alpha|\theta_i; \hat{T}_N^\alpha] = \binom{N-1}{\hat{T}_N^\alpha - 1}[\rho_i]^{\hat{T}_N^\alpha - 1}[1-\rho_i]^{N-\hat{T}_N^\alpha}, \ \forall i \in \{1,2,3\},$$

$$\mathbb{P}[piv_\beta|\theta_i; \hat{T}_N^\beta] = \binom{N-1}{\hat{T}_N^\beta - 1}[\rho_i]^{\hat{T}_N^\beta - 1}[1-\rho_i]^{N-\hat{T}_N^\beta}, \ \forall i \in \{1,2,3\},$$

$$\mathbb{P}[piv|\theta_i; \hat{\boldsymbol{T}}_{\boldsymbol{N}}, \mu_N] = \mu_N \cdot \mathbb{P}[piv_\alpha|\theta_i; \hat{T}_N^\alpha] + (1-\mu_N) \cdot \mathbb{P}[piv_\beta|\theta_i; \hat{T}_N^\beta], \ \forall i \in \{1,2,3\},$$

where $\hat{\boldsymbol{T}}_{\boldsymbol{N}} = (\hat{T}^\alpha, \hat{T}^\beta)$. The mechanism M_N must satisfy the incentive compatibility constraints under which the agents report truthfully,

$$\frac{q_3^0 \cdot \rho_3 \cdot \mathbb{P}[piv|\theta_3; \hat{\boldsymbol{T}}_{\boldsymbol{N}}, \mu_N] \cdot V_{ag}(\theta_3)}{-\sum_{i=1}^{2} q_i^0 \cdot \rho_i \cdot \mathbb{P}[piv|\theta_i; \hat{\boldsymbol{T}}_{\boldsymbol{N}}, \mu_N] \cdot V_{ag}(\theta_i)} \geq 1, \tag{1.A.35}$$

$$\frac{q_3^0 \cdot (1-\rho_3) \cdot \mathbb{P}[piv|\theta_3; \hat{\boldsymbol{T}}_{\boldsymbol{N}}, \mu_N] \cdot V_{ag}(\theta_3)}{-\sum_{i=1}^{2} q_i^0 \cdot (1-\rho_i) \cdot \mathbb{P}[piv|\theta_i; \hat{\boldsymbol{T}}_{\boldsymbol{N}}, \mu_N] \cdot V_{ag}(\theta_i)} \geq 1. \tag{1.A.36}$$

From (1.A.2),

$$\frac{\mathbb{P}[T;N|\theta_i]}{\mathbb{P}[T;N|\theta_{i'}]} = \exp\left\{N \cdot \left[KL\left(\frac{T}{N},\rho_{i'}x_h\right) - KL\left(\frac{T}{N},\rho_i x_h\right)\right]\right\}. \tag{1.A.37}$$

Note that for $i' > i$, if

$$t > \frac{\log\frac{1-\rho_i}{1-\rho_{i'}}}{\log\frac{\rho_i'}{\rho_i} + \log\frac{1-\rho_i}{1-\rho_{i'}}},$$

then

$$KL(t,\rho_{i'}) - KL(t,\rho_i) > 0.$$

While if

$$t < \frac{\log\frac{1-\rho_i}{1-\rho_{i'}}}{\log\frac{\rho_{i'}}{\rho_i} + \log\frac{1-\rho_i}{1-\rho_{i'}}},$$

then

$$KL(t,\rho_{i'}) - KL(t,\rho_i) < 0.$$

Hence, we can choose t_α and t_β such that

$$t_\alpha \in (\rho_1, \frac{\log\frac{1-\rho_1}{1-\rho_2}}{\log\frac{\rho_2}{\rho_1} + \log\frac{1-\rho_1}{1-\rho_2}}), \tag{1.A.38}$$

$$t_\beta \in (\frac{\log\frac{1-\rho_2}{1-\rho_3}}{\log\frac{\rho_3}{\rho_2} + \log\frac{1-\rho_2}{1-\rho_3}}, \rho_3), \tag{1.A.39}$$

and let $\hat{T}_N^\alpha$, $\hat{T}_N^\beta$ be the integers closest to Nt_α and Nt_β respectively. We have

$$\lim_{N\to\infty} \frac{\mathbb{P}[piv_\alpha|\theta_i;\hat{T}_N^\alpha]}{\mathbb{P}[piv_\alpha|\theta_1;\hat{T}_N^\alpha]} = 0; \forall i \in \{2,3\}, \tag{1.A.40}$$

$$\lim_{N\to\infty} \frac{\mathbb{P}[piv_\beta|\theta_i;\hat{T}_N^\beta]}{\mathbb{P}[piv_\beta|\theta_3;\hat{T}_N^\beta]} = 0; \forall i \in \{1,2\}. \tag{1.A.41}$$

By Sterling approximation and (1.A.2),

$$\frac{\mathbb{P}[piv_\alpha|\theta_i;\hat{T}_N^\alpha]}{\mathbb{P}[piv_\beta|\theta_j;\hat{T}_N^\beta]} = \exp\left\{N \cdot \left[KL\left(t_\beta,\rho_j\right) - KL\left(t_\alpha,\rho_i\right) + o(1)\right]\right\}, \ \forall i,j \in \{1,2,3\}. \tag{1.A.42}$$

Note that for each $i \in \{1,2,3\})$, the function $KL(t,\rho_i)$ strictly decreases with t when $t < \rho_i$ and strictly increases with t when $t > \rho_i$. We further have $KL(t,\rho_i) = 0$ if and only if $t = \rho_i$. We further choose t_α and t_β such that

$$KL(t_\alpha,\rho_1) > KL(t_\beta,\rho_3),$$

and choose $\mu_N \in (0,1)$ such that

$$\frac{\mu_N}{1-\mu_N} \cdot \frac{\mathbb{P}[piv_\alpha|\theta_1;\hat{T}_N^\alpha]}{\mathbb{P}[piv_\beta|\theta_3;\hat{T}_N^\beta]} \cdot \frac{-V_{ag}(\theta_1)}{V_{ag}(\theta_3)} = 1. \tag{1.A.43}$$

From (1.A.42), we can see that μ_N must exist when N is large enough and

$$\lim_{N\to\infty} \mu_N = 0. \tag{1.A.44}$$

From (1.A.41) and (1.A.43),

$$\lim_{N\to\infty} \frac{\mu_N}{1-\mu_N} \cdot \frac{\mathbb{P}[piv_\alpha|\theta_1;\hat{T}_N^\alpha]}{\mathbb{P}[piv_\beta|\theta_1;\hat{T}_N^\beta]} = \infty.$$

Hence,

$$\lim_{N\to\infty} \frac{\mu_N \cdot \mathbb{P}[piv_\alpha|\theta_1;\hat{T}_N^\alpha]}{\mathbb{P}[piv|\theta_1;\hat{\boldsymbol{T}}_{\boldsymbol{N}},\mu_N]} = 1.$$

Similarly,

$$\lim_{N\to\infty} \frac{(1-\mu_N) \cdot \mathbb{P}[piv_\beta|\theta_3;\hat{T}_N^\beta]}{\mathbb{P}[piv|\theta_3;\hat{\boldsymbol{T}}_{\boldsymbol{N}},\mu_N]} = 1.$$

Furthermore, from (1.A.41), (1.A.42) and (1.A.43),

$$\lim_{N\to\infty} \frac{\mathbb{P}[piv|\theta_2;\hat{\boldsymbol{T}}_{\boldsymbol{N}},\mu_N]}{\mathbb{P}[piv|\theta_1;\hat{\boldsymbol{T}}_{\boldsymbol{N}},\mu_N]} = 0,$$

$$\lim_{N\to\infty} \frac{\mathbb{P}[piv|\theta_2;\hat{\boldsymbol{T}}_{\boldsymbol{N}},\mu_N]}{\mathbb{P}[piv|\theta_3;\hat{\boldsymbol{T}}_{\boldsymbol{N}},\mu_N]} = 0.$$

Therefore, when N is large,

$$\frac{q_3^0 \cdot \mathbb{P}[piv|\theta_3;\hat{\boldsymbol{T}}_{\boldsymbol{N}},\mu_N] \cdot V_{ag}(\theta_3)}{-\sum_{i=1}^2 q_i^0 \cdot \mathbb{P}[piv|\theta_i;\hat{\boldsymbol{T}}_{\boldsymbol{N}},\mu_N] \cdot V_{ag}(\theta_i)} \approx \frac{\mu_N}{1-\mu_N} \cdot \frac{\mathbb{P}[piv_\alpha|\theta_1;\hat{T}_N^\alpha]}{\mathbb{P}[piv_\beta|\theta_3;\hat{T}_N^\beta]} \cdot \frac{-V_{ag}(\theta_1)}{V_{ag}(\theta_3)} = 1,$$

by which (1.A.35) and (1.A.36) are satisfied and we hence construct an incentive-compatible mechanism. By (1.A.38), (1.A.39), (1.A.44), and the law of large numbers, we can see that the principal can approach his first-best outcome as $N \to \infty$.

References

Austen-Smith, David. 1990. "Information transmission in debate." *American Journal of Political Science* 34 (1): 124–52. [37]

Austen-Smith, David. 1993. "Interested experts and policy advice: Multiple referrals under open rule." *Games and Economic Behavior* 5 (1): 3–43. [37]

Battaglini, Marco. 2017. "Public protests and policy making." *Quarterly Journal of Economics* 132 (1): 485–549. [3, 4, 6, 14, 28, 36]

Chen, Ying, Navin Kartik, and Joel Sobel. 2008. "Selecting cheap-talk equilibria." *Econometrica* 76 (1): 117–36. [37]

Chernoff, Herman. 1952. "A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations." *Annals of Mathematical Statistics* 23 (4): 493–507. [6]

Cho, In-Koo, and Joel Sobel. 1990. "Strategic stability and uniqueness in signaling games." *Journal of Economic Theory* 50 (2): 381–413. [37]

Crawford, Vincent P, and Joel Sobel. 1982. "Strategic information transmission." *Econometrica* 50 (6): 1431–51. [6, 27]

Currarini, Sergio, Giovanni Ursino, and AKS Chand. 2020. "Strategic transmission of correlated information." *Economic Journal* 130 (631): 2175–206. [37]

Duggan, John, and César Martinelli. 2001. "A Bayesian model of voting in juries." *Games and Economic Behavior* 37 (2): 259–94. [13, 34, 36]

Ekmekci, Mehmet, and Stephan Lauermann. 2022. "Informal elections with dispersed information." Working paper. [3, 14, 36]

Feddersen, Timothy, and Wolfgang Pesendorfer. 1997. "Voting behavior and information aggregation in elections with private information." *Econometrica* 65 (5): 1029–58. [13, 34, 36]

Feddersen, Timothy, and Wolfgang Pesendorfer. 1998. "Convicting the innocent: The inferiority of unanimous jury verdicts under strategic voting." *American Political Science Review* 92 (1): 23–35. [13, 34, 36]

Frick, Mira, Ryota Iijima, and Yuhta Ishii. 2021a. "Belief convergence under misspecified learning: A martingale approach." Working paper. [32]

Frick, Mira, Ryota Iijima, and Yuhta Ishii. 2021b. "Welfare comparisons for biased learning." Working paper. [37]

Galeotti, Andrea, Christian Ghiglino, and Francesco Squintani. 2013. "Strategic information transmission networks." *Journal of Economic Theory* 148 (5): 1751–69. [37]

Gerardi, Dino, Richard McLean, and Andrew Postlewaite. 2009. "Aggregation of expert opinions." *Games and Economic Behavior* 65 (2): 339–71. [37]

Gordon, Sidartha, Navin Kartik, Melody Pei-yu Lo, Wojciech Olszewski, and Joel Sobel. 2021. "Effective communication in cheap talk games." Working paper. [37]

Hagenbach, Jeanne, and Frédéric Koessler. 2010. "Strategic communication networks." *Review of Economic Studies* 77 (3): 1072–99. [37]

Ivanov, Maxim. 2010. "Communication via a strategic mediator." *Journal of Economic Theory* 145 (2): 869–84. [37]

Kolotilin, Anton, and Hongyi Li. 2021. "Relational communication." *Theoretical Economics* 16 (4): 1391–430. [37]

Krishna, Vijay, and John Morgan. 2001. "A model of expertise." *Quarterly Journal of Economics* 116 (2): 747–75. [37]

Ladha, Krishna K. 1992. "The Condorcet jury theorem, free speech, and correlated votes." *American Journal of Political Science* 36 (3): 617–34. [13]

Levit, Doron, and Nadya Malenko. 2011. "Nonbinding voting for shareholder proposals." *Journal of Finance* 66 (5): 1579–614. [3, 4, 14, 36]

Morgan, John, and Phillip C Stocken. 2008. "Information aggregation in polls." *American Economic Review* 98 (3): 864–96. [3, 36, 37]

Moscarini, Giuseppe, and Lones Smith. 2002. "The law of large demand for information." *Econometrica* 70 (6): 2351–66. [37]

Myerson, Roger B. 1998. "Extended Poisson games and the Condorcet jury theorem." *Games and Economic Behavior* 25 (1): 111–31. [13, 34, 36]

Razin, Ronny. 2003. "Signaling and election motivations in a voting model with common values and responsive candidates." *Econometrica* 71 (4): 1083–119. [36]

Vida, Peter, Takakazu Honryo, and Helmuts Azacis. 2022. "Strong forward induction in monotonic multi-sender signaling games." Working paper. [37]

Wolinsky, Asher. 2002. "Eliciting information from multiple experts." *Games and Economic Behavior* 41 (1): 141–60. [3, 36]

Chapter 2

Information Aggregation in Collective Experimentation

2.1 Introduction

Attention is often focused on the fight to implement reform in policies. And yet, in order to evaluate the overall suitability of reform, it is important to study the divergent effects that it has on a population. Some may benefit enormously while others are worse off than before, and the full effects of reform may need time to arrive. This disparity of results in turn complicates the information available to a population to measure its suitability. For example, after the UK joined the European Union, British citizens deliberated whether the gradual changes implemented by this reform primarily caused them to gain benefits or suffer hardships. The answer was not a simple one, as it varied from sector to sector and from time to time. Each person had a piece of the information but no one had any easy access to the whole picture of the situation that would allow them to judge whether joining the EU was to the advantage of the majority. On 23 June 2016, British citizens voted by referendum to leave the EU. Did this election aggregate people's private information? Was BREXIT the utilitarian optimum?

This paper analyzes a collective experimentation model in which voters gradually learn their payoffs, which are divergent among them. Furthermore, their payoffs depend on an unknown state of the world. Hence, experimentation generates information concerning the unknown state, which is dispersed among voters. We are interested in how strategic voting shapes incentives for experimentation, and more importantly, whether elections can aggregate and utilize voters' private information concerning the unknown state.

We study a two-armed exponential bandit model with N voters. At each instant, voters vote between a safe action and a risky action. The risky action is chosen if the ratio of votes for it exceeds a given number $k \in (0, 1]$. If not, then the safe action is chosen irreversibly. The safe action yields a constant, homogeneous payoff to every voter, while the risky action yields payoffs depending on voters' types, which are initially unobservable to all voters. For each voter, if her type is bad, then the risky action always pays her nothing. If her type is good, then the risky action pays her lump-sum payoffs at random times corresponding to the jumping times of a Poisson process. When a voter receives a lump-sum, she becomes a sure winner, that is, she is informed that her type is good. By contrast, unsure voters are those who have not received any lump-sum. They become more pessimistic as experimentation on the risky action goes on. We further assume that each voter can only observe her own payoff stream.

Voters' payoffs are correlated. Before the game starts, nature chooses a state $\omega \in \{H, L\}$ randomly. Voters are uncertain about the state. After choosing the state, nature chooses the type of each voter randomly and independently. Each voter's type is more likely to be good in state H than in state L. Hence, knowing that there are more sure winners makes unsure voters believe that their types are more likely to be good, creating additional incentives for experimentation.

We focus on the symmetric pure-strategy equilibrium and further assume that each voter will always vote for the risky action after she becomes a sure winner. Hence, unsure voters use the same cut-off strategy such that they vote for the risky action before a cut-off time $\hat{t}$ and vote for the safe action after that in the equilibrium. We further require that unsure voters are indifferent between the risky action and safe action at the cut-off time $\hat{t}$ to rule out equilibria in which each unsure voter votes for the risky action when she strictly prefers the safe action but her vote cannot change the election outcome.

Strulovici (2010) has focused on a similar setting in which voters' payoffs are publicly observed and independent. He shows that incentives for experimentation are always weaker compared to the case of a single decision-maker since the control power over future experimentation is shared. Strategic voting shapes incentives for experimentation differently in our paper where voters' payoffs are privately observed and correlated. Each unsure voter makes decisions at the cut-off time $\hat{t}$ conditional on the event that her vote can change the election outcome, that is, conditional on being pivotal. She updates her belief about the realized state conditional on the event that there are exactly $kN - 1$ sure winners[1]. Thus, strategic voting conveys information and affects voters' incentives for experimentation.

We analyze the limit properties of the equilibrium cut-off time when the number of voters goes to infinity. We find that the limit cut-off time is increasing in k, that is, a stricter voting rule leads to more experimentation. However, the limit cut-off time is bounded above by the stopping time chosen by a myopic decision-maker who is sure that the realized state is H. For the intuition, when the number

1. Assume kN is an integer.

of voters is large, each unsure voter updates her belief conditional on the event that she is pivotal and only can be pivotal at the cut-off time, acting myopically since individual control over future decisions becomes infinitely diluted. Unsure voters are more optimistic that the realized state is H and their types are good when k is higher since they infer that there are more sure winners conditional on being pivotal. Therefore, a higher k leads to more experimentation. However, when unsure voters are sure that the realized state is H, a higher k cannot increase their belief of being good type. Hence, the limit cut-off time is bounded above.

When experimenting with the risky action, voters learn gradually about their own types. Hence, they gain private information about the realized state. We are interested in whether elections can aggregate the voters' information generated by experimentation and pick the most suitable action in each state. Specifically, we assume that the risky action has a higher expected flow payoff in state H while the safe action has a higher expected flow payoff in state L. We are wondering if the risky action is always chosen after the limit cut-off time in state H and the safe action is always chosen after the limit cut-off time in state S when the number of voters goes to infinity. In contrast, there is no uncertainty about which action has a higher expected flow payoff in Strulovici (2010). The one with a higher expected flow payoff is always chosen when the number of voters goes to infinity.

We show that information aggregation obtains if k is below some threshold, that is, if the ratio of votes required for the risky action R is low. Hence, we argue that the voting rule should be biased toward experimentation. For intuition, note that experimentation reveals the true state but brings heterogeneity among voters in their beliefs of being good type. From the previous result, a higher k leads to more experimentation. When k is small, the heterogeneity in beliefs among voters is small. Sure winners prefer the risky action in both states while unsure voters prefer the risky action in state H and the safe action in state L. We prove that the vote share for the risky action must be bigger than k in state H and lower than k in state L at the cut-off time when the number of voters goes to infinity. If not, then each unsure voter would believe they are in the state where the vote share for the risky action is closer to k for sure conditional on being pivotal, and she cannot be indifferent between the two actions at the cut-off time. The proof in this case is similar to the one for information aggregation in the static voting model (Feddersen and Pesendorfer (1997), Duggan and Martinelli (2001)), in which strategic voters make inferences conditional on being pivotal and try to match different actions to different states. When k is large, the heterogeneity in beliefs among voters is large. Sure winners prefer the risky action in both states while unsure voters prefer the safe action in both states. There are unsure voters who should vote for the risky action if they know their types but vote for the safe action due to a pessimistic belief. Hence, the heterogeneity in beliefs generates a bias towards the safe action. Information is not aggregated. This case relates to

Fernandez and Rodrik (1991), which explains the status-quo bias in the presence of asymmetric uncertainty about the payoffs from a new reform. They consider a static voting model in which both the voting rule and asymmetric uncertainty are exogenous. In the present paper, asymmetric uncertainty is endogenized by the voting rule.

This paper contributes to the literature on experimentation with multiple agents started by Bolton and Harris (1999) and Keller, Rady, and Cripps (2005). Most of the literature in this field assumes that the payoff of each agent is publicly observed. Hence, an agent can directly learn from others if their payoffs are correlated. By contrast, we assume the payoff of each voter is privately observed. Each voter indirectly learns from others by conditioning on being pivotal. Hence, different voting rules generate different incentives for experimentation. Similarly, Halac, Kartik, and Liu (2017) study contests for experimentation in which each agent indirectly learns from others by conditioning on the event that the contest is not stopped by the principal. They show that the principal can generate more incentives for experimentation by committing to share the prize and stop the contest following a certain number of successes instead of only awarding the first success.

This paper is related to the literature on information aggregation in elections with strategic voters started by Austen-Smith and Banks (1996), Feddersen and Pesendorfer (1997), and Feddersen and Pesendorfer (1998). Most of the literature in this field analyzes a static model in which the preferences of voters are exogenous. Voters receive information purely about the unknown state that affects all voters' payoffs in the same direction. We instead focus on a dynamic model in which voters receive more precise information about their own preferences and the unknown state as experimentation evolves. We argue that more precise information might reduce the total welfare if it enlarges the heterogeneity among voters asymmetrically.

This paper shares common features with Murto and Välimäki (2011). They analyze information aggregation in a stopping game with uncertain payoffs that are correlated among players. In their model, players make decisions about when to exit the game individually while in our model voters choose between two actions collectively. Moldovanu and Rosar (2021) consider a similar setting in which voters jointly decide between a reversible option and an irreversible one. They also argue that the voting rule should be biased toward the reversible option. Otherwise, the coordination failure might diminish the option value from the reversible option.

This paper proceeds as follows: Section 2.2 describes the model. Section 2.3 characterizes the equilibrium. Section 2.4 analyzes the limit case in which the number of voters goes to infinity and presents our main results about information aggregation. Section 2.5 concludes.

2.2 Model

2.2.1 Model Setting

We study an exponential bandit model in continuous time with $t \in [0, \infty)$. Payoffs are discounted at rate r. There is a set of voters, denoted by $\{1, ..., N\}$ with $N \geq 1$. Voters vote continuously on two actions, the safe action S and the risky action R. Let $k \in (0, 1]$ be a fixed number. At each time $t \in [0, \infty)$, the risky action R is chosen if the number of votes for R is larger than or equal to $k \cdot N$. If not, then the game ends and the safe action S is chosen irreversibly.[2] At each time, the aggregate numbers of votes for each alternative are observed. Assume that kN is an integer.[3]

Before the game starts, nature chooses a state randomly from two alternatives $\{H, L\}$. Voters are uncertain about the state. They hold a common prior belief q_0 that the state is H. After choosing the state, nature chooses the type of each voter randomly and independently. For each voter, her type is either good or bad. If the state is H, then the probability of being good type is ρ_H, while if the state is L, then the probability of being good type is ρ_L. Each voter is more likely to be of good type in state H than in state L ($\rho_H > \rho_L$). All types are initially unobservable to all voters.

If the safe action S is chosen, then it yields a flow s per unit of time to all voters. If the risky action R is chosen, then for each voter $i \in \{1, ..., N\}$, the payoff depends on her type. If her type is bad, then R always pays her 0. If her type is good, then R pays her lump-sum payoffs at random times corresponding to the jumping times of a Poisson process with constant intensity $\lambda > 0$. The arrivals of lump-sums are independent among individuals. The magnitude of these lump-sums equals z. We denote the expected payoff per unit of time from R when the voter's type is good by $g = \lambda z$. Assume that

$$\rho_H g > s > \rho_L g > 0. \tag{2.1}$$

Based on the prior belief, the risky action R has a higher expected flow payoff in state H while the safe action S has a higher expected flow payoff in state L. Payoffs are privately observed. Each voter can only observe her own payoff stream.

At each time $t \in [0, \infty)$, we can divide voters into two groups. One is ***sure winners*** who received lump-sums before t. They are sure that their types are good. The other is ***unsure voters*** who have not received any lump-sum yet. They share the same belief that their types are good.

2. We consider the case in which it is costly to restart the risky policy.

3. All results can be extended to the general case in which R is chosen if and only if at least $\lceil kN \rceil$ voters vote for R.

2.2.2 Equilibrium

We say a voter follows a ***cut-off strategy*** $\hat{\boldsymbol{t}}$ if she votes for R for $t < \hat{t}$ and votes for S for $t \geq \hat{t}$. We focus on Bayesian Nash equilibrium in which all unsure voters follow an identical cut-off strategy.[4]

Definition 2.1. A Bayesian Nash equilibrium is ***simple*** if

(1) sure winners always vote for R,

(2) unsure voters all use a cut-off strategy $\hat{t}_{k,N}$ with $\hat{t}_{k,N} > 0$,

(3) at time $\hat{t}_{k,N}$, each unsure voter must be indifferent between voting for R and voting for S.

A simple equilibrium is solely characterized by the cut-off $\hat{t}_{k,N}$. All voters vote for R at $t < \hat{t}_{k,N}$. Hence, the risky action R is chosen at $t < \hat{t}_{k,N}$. At the cut-off time $\hat{t}_{k,N}$, if there are at least kN sure winners, then R is chosen forever. If not, then S is chosen forever.

Requirement (iii) of Definition 2.1 is a refinement. Suppose unsure voters strictly prefer to vote for S at time $\hat{t}_{k,N}$. Consider the case in which unsure voters use a mixed strategy mixing the cut-off strategy $\hat{t}_{k,N}$ with probability $1-\epsilon$ and the cut-off strategy $\hat{t}_{k,N} - \delta$ with probability ϵ for some $\delta > 0$. By choosing a small δ, we can see that each unsure voter strictly prefers to vote for S at time $\hat{t}_{k,N} - \delta$, no matter how small ϵ is.

In the simple equilibrium, consider an unsure voter at the cut-off time $\hat{t}_{k,N}$. If there are more than $kN-1$ sure winners, then R is chosen forever whenever this unsure voter votes for R or S. If there are fewer than $kN-1$ sure winners, then S is chosen forever whenever she votes for R or S. Hence, this voter's vote can change the election outcome only if there are exactly $kN-1$ sure winners. Therefore, at the cut-off time $\hat{t}_{k,N}$, each unsure voter makes decisions conditional on the event that there are exactly $kN-1$ sure winners, that is, conditional on being ***pivotal*** for the election outcome and having ***full control*** of experimentation. Hence, in a simple equilibrium, even though voters do not observe the payoffs of others, they can still learn something about the aggregate state conditional on being pivotal. Thus, strategic voting conveys information and affects voters' incentives for experimentation.

2.2.3 Beliefs

Consider the case in which all voters always vote for R before some time $t > 0$, that is, the risky action R is always chosen before t.

4. Since the safe action is irreversible, we actually consider the symmetric pure-strategy equilibrium in which each voter always votes for R after becoming the sure winner.

Denote the probability that the type of an unsure voter is good at time t conditional on state H by $p(t|H)$, and denote the probability that the type of an unsure voter is good at time t conditional on state L by $p(t|L)$:

$$p(t|H) = \frac{\rho_H e^{-\lambda t}}{\rho_H e^{-\lambda t} + 1 - \rho_H}, \tag{2.2}$$

$$p(t|L) = \frac{\rho_L e^{-\lambda t}}{\rho_L e^{-\lambda t} + 1 - \rho_L}. \tag{2.3}$$

Note that the probability that a voter of good type has not yet received any lump-sum yet is $e^{-\lambda t}$ in each state.

Denote the probability that the state is H conditional on the event that there are K sure winners at t by $q(K,t)$. In state ω, the probability that a voter receives lump-sums before t is $\rho_\omega(1-e^{-\lambda t})$. Hence,

$$\frac{q(K,t)}{1-q(K,t)} = \underbrace{\frac{q_0}{1-q_0}}_{\text{prior}} \underbrace{\left[\frac{\rho_H(1-e^{-\lambda t})}{\rho_L(1-e^{-\lambda t})}\right]^K}_{K \text{ sure winners}} \underbrace{\left[\frac{1-\rho_H(1-e^{-\lambda t})}{1-\rho_L(1-e^{-\lambda t})}\right]^{N-K}}_{N-K \text{ unsure voters}}. \tag{2.4}$$

Finally, denote the belief of an unsure voter that her type is good conditional on the event that there are K sure winners at time t by $p(K,t)$,

$$p(K,t) = \underbrace{q(K,t)p(t|H)}_{H \text{ state and good type}} + \underbrace{(1-q(K,t))p(t|L)}_{L \text{ state and good type}}. \tag{2.5}$$

Notice that $p(t|H)$ and $p(t|L)$ are strictly decreasing in t. Also, notice that $q(K,t)$ and $p(K,t)$ are strictly increasing in K and strictly decreasing in t.

2.3 Characterization

We calculate the cut-off time $\hat{t}_{k,N}$ in a simple equilibrium. At $\hat{t}_{k,N}$, each unsure voter makes decisions conditional on being pivotal, that is, conditional on the event that there are $kN-1$ sure winners. She believes that her type is good with probability $p(kN-1,\hat{t}_{k,N})$. Since unsure voters are indifferent between voting for R and voting for S at $\hat{t}_{k,N}$, the cut-off time is supposed be a solution to the equation,

$$s = \underbrace{p(kN-1,\hat{t})g}_{\text{flow payoff from } R} + \underbrace{p(kN-1,\hat{t})\lambda\left(\frac{g}{r}-\frac{s}{r}\right)}_{\text{jump when } i \text{ receives a lump-sum}} \underbrace{+(N-kN)p(kN-1,\hat{t})\lambda\left[\frac{p(kN,\hat{t})g}{r}-\frac{s}{r}\right]}_{\text{jump when others receive a lump-sum}}. \tag{2.6}$$

Consider an unsure voter i at the cut-off time $\hat{t}_{k,N}$. Conditional on being pivotal, this voter has full control of experimentation. If she votes for R (S), then R (S) is chosen for the next instant. The left side of equation (2.6) is the flow payoff by choosing S. The right side contains the flow payoff of choosing R for the next instant and the jumps in the discounted payoff when voter i receives a lump-sum and when other unsure voters receive lump-sums. When voter i receives a lump-sum in the next instant with probability $p(kN-1,\hat{t})\lambda dt$, she becomes a sure winner and R is chosen forever. Her expected payoff jumps from $\frac{s}{r}$ to $\frac{g}{r}$. When one of the other unsure voters receives a lump-sum in the next instant with probability $(N-kN)p(kN-1,\hat{t})\lambda dt$, the risky action R is chosen forever and i remains as an unsure voter with belief $p(kN,\hat{t}_{k,N})$. Her expected payoff jumps from $\frac{s}{r}$ to $\frac{p(kN,\hat{t}_{k,N})g}{r}$

Proposition 2.1. *For each $k \in (0,1]$, there exists $N_k > 0$ such that for each $N > N_k$, a unique simple equilibrium exists.*

2.4 Large Elections

In this section, we analyze the limiting properties of the sequence of simple equilibria with N voters with $N \to \infty$.

2.4.1 Limit Cut-off

We characterize the limit of the sequence of cut-offs $\{\hat{t}_{k,N}\}$ as $N \to \infty$. Pick $\bar{t}$ such that $p(\bar{t}|H) = \frac{s}{g}$. Note that $\bar{t}$ is the time when a single myopic decision-maker stops if she knows the state is H but has not received any lump-sum.

Proposition 2.2. *For each $k \in (0,1]$, there exists $\hat{t}_k \in (0,\bar{t}\,]$ such that*

$$\lim_{N\to\infty} \hat{t}_{k,N} = \hat{t}_k.$$

There exists $\bar{k} \in (0,1)$ such that $\hat{t}_k$ is strictly increasing in k when $k \leq \bar{k}$ with $\lim_{k\to 0} \hat{t}_k = 0$ and $\hat{t}_{\bar{k}} = \bar{t}$. When $k > \bar{k}$, the limit cut-off $\hat{t}_k$ is equal to $\bar{t}$.

We illustrate Proposition 2.2 by Figure 2.1. When $k < \bar{k}$, the limit cut-off $\hat{t}_k$ is increasing in k. Voters experiment more under a stricter voting rule. However, when $k \geq \bar{k}$, the limit cut-off $\hat{t}_k$ equals the myopic cut-off $\bar{t}$. Each unsure voter acts like a single myopic decision-maker who believes the realized state is H. In addition, the limit cut-off $\hat{t}_k$ is continuous in k and bounded above by $\bar{t}$ for each $k \in (0,1]$.

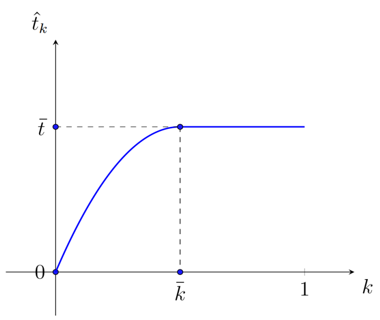

Figure 2.1. Limit cut-off

When the number of voters is large and other unsure voters use the cut-off strategy $\hat{t}$, suppose that voter i is an unsure voter and pivotal at $\hat{t}$. If she votes for R during the next instant $[\hat{t}, \hat{t} + dt)$ instead of S, then R is chosen forever if an unsure voter receives a lump-sum in the next instant. However, it is rare that she is the one who receives a lump-sum and makes the risky action R be chosen forever. It is almost certain that other unsure voters receive lump-sums and voter i remains an unsure voter with no control power. Thus, when the number of voters is large, voter i updates her belief about the realized state conditional on the event that she is pivotal and can only be pivotal at time $\hat{t}$, or say, conditional on the event that she has full control of experimentation at $\hat{t}$ and will lose the control forever in the next instant. Voter i acts like a single myopic decision-maker, choosing between R and S forever.

Since unsure voters vote conditional on being pivotal, that is, conditional on the event that there are $kN - 1$ sure winners, they are more optimistic that the realized state is H and their types are good when k is higher. Hence, a higher k leads to a higher limit cut-off $\hat{t}_k$. However, this effect is limited. Even if unsure voters are certain that the realized state is H, they vote for S after the time $\bar{t}$ since they behave myopically. Hence, the limit cut-off $\hat{t}_k$ is bounded above by $\bar{t}$ for each $k \in (0, 1]$.

2.4.2 Information Aggregation

When experimenting with the risky action R, voters learn gradually about their own types, hence, they gain private information about the aggregate state. We are interested in whether information aggregation obtains when the number of voters goes to infinity. From (2.1), based on the prior belief, the risky action R has a higher expected flow payoff in state H while the safe action S has a higher expected flow payoff in state L. Hence, according to the law of large numbers, when $N \to \infty$, under the optimal decision made by a utilitarian social planner who knows the state and the type of each voter, the probability that R is chosen in state H goes to 1, and the probability that S is chosen in state L goes to 1. We now formally define information aggregation for every sequence of strategy profiles with $N \to \infty$.

Definition 2.2. A sequence of strategy profiles ***aggregates information*** if there is a $\hat{t} > 0$ and for each $\epsilon > 0$, there is an N_ϵ such that for $N > N_\epsilon$, the following holds:

(1) the event that R is chosen at each $t \in [\hat{t}, \infty)$ in state H happens with probability greater than $1 - \epsilon$,

(2) the event that S is chosen at each $t \in [\hat{t}, \infty)$ in state L happens with probability greater than $1 - \epsilon$.

In the simple equilibrium, denote the numbers of sure winners at cut-off time $\hat{t}_{k,N}$ in state H and state L by $H(\hat{t}_{k,N})$ and $L(\hat{t}_{k,N})$. The sequence of simple equilibria aggregates information if

$$\lim_{N\to\infty} Pr(H(\hat{t}_{k,N}) > kN) = 1,$$
$$\lim_{N\to\infty} Pr(L(\hat{t}_{k,N}) < kN) = 1.$$

We now provide the condition under which the sequence of simple equilibria aggregates information.

Proposition 2.3. *For each $k \in (0, \frac{\rho_H g - s}{g-s})$, the sequence of simple equilibria aggregates information, that is,*

$$\lim_{N\to\infty} Pr(H(\hat{t}_{k,N}) > kN) = 1,$$
$$\lim_{N\to\infty} Pr(L(\hat{t}_{k,N}) < kN) = 1.$$

For each $k \in [\frac{\rho_H g - s}{g-s}, 1]$, information is not aggregated with

$$\lim_{N\to\infty} Pr(H(\hat{t}_{k,N}) > kN) < 1.$$

Proposition 2.3 shows that the sequence of simple equilibria aggregates information if k is small, that is, if the ratio of votes required for the risky action R is low. We hence argue that the voting rule should be biased toward experimentation.

For intuition, experimentation brings information about the aggregate state. However, during experimentation, voters' beliefs evolve and become heterogeneous. From Proposition 2.2, a higher k leads to more experimentation. When $k < \bar{k}$, the heterogeneity in beliefs among voters is small. Conditional on being pivotal at $\hat{t}_k$, sure winners prefer R in both states while unsure voters prefer R in state H and S in state L. We show that strategic voting successfully conveys information through pivotal reasoning and leads to information aggregation. When $k > \bar{k}$, the heterogeneity in beliefs among voters is large. Conditional on being pivotal at $\hat{t}_k$, sure winners prefer R in both states while unsure voters prefer S in both states. There are unsure voters who should vote for R if they know their types but vote for S due to a pessimistic belief. The heterogeneity in beliefs generates a bias towards S and leads to the failure of information aggregation when $k > \frac{\rho_H g - s}{g-s}$.

We now sketch the proof for Proposition 2.3. By the law of large numbers and Proposition 2.2, the ratios of votes for R at the cut-off time $\hat{t}_{k,N}$ in both states converge,

$$\frac{H(\hat{t}_{k,N})}{N} \xrightarrow{p} \rho_H(1 - e^{-\lambda \hat{t}_k}),$$

$$\frac{L(\hat{t}_{k,N})}{N} \xrightarrow{p} \rho_L(1-e^{-\lambda\hat{t}_k}).$$

The sequence of simple equilibria aggregates information if

$$\rho_H(1-e^{-\lambda\hat{t}_k}) \;>\; k \;>\; \rho_L(1-e^{-\lambda\hat{t}_k}).$$

We draw both limit ratios of votes $\rho_H(1-e^{-\lambda\hat{t}_k})$ and $\rho_L(1-e^{-\lambda\hat{t}_k})$ as functions of k in Figure 2.2. By Proposition 2.2, they are both strictly increasing in k when $k<\bar{k}$ and constant when $k \geq \bar{k}$. We also draw the 45 degree line in Figure 2.2 (the blue line). For each $k \in (0,1]$, the sequence of simple equilibria aggregates information if at this k, the red line is above the blue line while the green line is below the blue line.

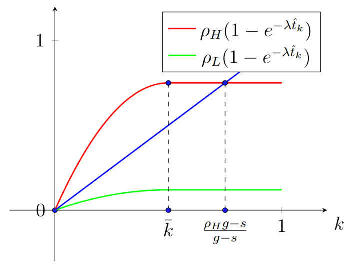

Figure 2.2. Limit ratios of votes

<u>Case 1: $k<\bar{k}$</u>

In this case, we have $\hat{t}_k < \bar{t}$. Hence, conditional on being pivotal at $\hat{t}_k$, unsure voters prefer R in state H and S in state L since they act myopically as $N \to \infty$. The proof in this case is similar to the one for information aggregation in the static voting model (Feddersen and Pesendorfer (1997), Duggan and Martinelli (2001)), in which strategic voters make inferences conditional on being pivotal and try to match different actions to different states. Similar to these studies, we demonstrate that strategic voting successfully conveys information through pivotal reasoning and leads to information aggregation. We show that

$$\rho_H(1-e^{-\lambda\hat{t}_k}) \;>\; k \;>\; \rho_L(1-e^{-\lambda\hat{t}_k}),\;\; \forall k \;<\; \bar{k}.$$

Otherwise, if

$$k \;>\; \rho_H(1-e^{-\lambda\hat{t}_k}) \;>\; \rho_L(1-e^{-\lambda\hat{t}_k}),$$

unsure voters realize that the state must be H conditional on being pivotal at $\hat{t}_k$, while if

$$\rho_H(1-e^{-\lambda\hat{t}_k}) \;>\; \rho_L(1-e^{-\lambda\hat{t}_k}) \;>\; k,$$

unsure voters realize that the state must be L conditional on being pivotal at $\hat{t}_k$. In both cases, unsure voters cannot be indifferent at $\hat{t}_k$.

Case 2: $k \geq \bar{k}$

In this case, since the limit cut-off $\hat{t}_k$ equals $\bar{t}$ by Proposition 2.2, the vote share for R at $\hat{t}_k$ stays constant for all $k \geq \bar{k}$,

$$\rho_H(1-e^{-\lambda \hat{t}_k}) = \rho_H(1-e^{-\lambda \bar{t}}) = \frac{\rho_H g - s}{g-s}.$$

Hence, information is not aggregated in state H for each $k \in [\frac{\rho_H g - s}{g-s}, 1)$.

Note that when $k \geq \bar{k}$, each unsure voter is sure that the realized state is H conditional on being pivotal at $\bar{t}$ when $N \to \infty$,

$$\lim_{N\to\infty} \frac{q(kN-1,\bar{t})}{1-q(kN-1,\bar{t})} = \infty, \quad \forall k \geq \bar{k}.$$

However, each unsure vote acts myopically and votes for S at $\bar{t}$ even if she is sure that they are in state H since she holds a pessimistic belief about her type.

The inefficiency in this case is generated by asymmetric uncertainty. Assume that voters know the realized state is H. Consider the following two cases. (i) If there is no uncertainty, that is, each voter knows her own type, then R is chosen if $k < \rho_H$ due to the law of large numbers. (ii) If the uncertainty is symmetric, that is, each voter holds the same belief that her type is good with probability ρ_H, then R is chosen if $k \leq 1$. However, in the simple equilibrium, there exists uncertainty since voters act myopically and the cut-off time $\hat{t}_k$ is bounded above by $\bar{t}$. In addition, the uncertainty is asymmetric. Sure winners are sure that their types are good while unsure voters hold a pessimistic belief about their types. There are unsure voters whose types are good. They would vote for R if they knew their types, but they vote for S even if they are sure the realized state is H. Hence, there is a bias towards S. The safe action S is chosen in state H when $k > \frac{\rho_H g - s}{g-s}$ with $\frac{\rho_H g - s}{g-s} < \rho_H$.

This case relates to Fernandez and Rodrik (1991), which explains the status-quo bias in the presence of asymmetric uncertainty about the payoffs from a new reform. They consider a static voting model in which both the voting rule and asymmetric uncertainty are exogenous. In the present paper, asymmetric uncertainty is endogenized by the voting rule. A stricter voting rule leads to more experimentation and generates larger heterogeneity in beliefs among voters.

From Proposition 2.3, we can approximate the first-best outcome by choosing an arbitrarily small k since $\lim_{k\to 0} \hat{t}_k = 0$. Information aggregation obtains almost immediately. However, information aggregation fails when voters follow the unanimity rule over the safe action S under which the risky action R is chosen if at least one voter votes for it.[5]

5. If R is always chosen until some time t, then the probability that there is no sure winner in state $\omega \in \{H, L\}$ is $[1-\rho_\omega(1-e^{-\lambda t})]^N$. We obtain information aggregation if there exists $\hat{t}(N)$ such that $\lim_{N\to\infty}[1-\rho_H(1-e^{-\lambda \hat{t}(N)})]^N = 0$ and $\lim_{N\to\infty}[1-\rho_L(1-e^{-\lambda \hat{t}(N)})]^N = 1$. However, those two equations cannot happen at the same time.

2.5 Concluding Remarks

This paper studies a dynamic voting model in which voters jointly decide whether or not to experiment with a risky action. We show how strategic voting affects voters' incentives for experimentation and provide conditions under which information aggregation obtains.

We assume that the decision to switch the safe action is irreversible, which corresponds to the case in which it is costly to restart the risky policy. If the decision is reversible, there exist multiple symmetric pure-strategy equilibria including the simple equilibrium and the Markov equilibrium in undominated strategies analyzed by Strulovici (2010). However, since the votes of unsure voters matter only if the number of sure winners is less than kN and unsure voters' belief of being good type is increasing in the number of sure winners, it is reasonable to focus on the equilibria in which unsure voters always vote for the safe action after the cut-off time characterized in the simple equilibrium.[6] Hence, if the decision to switch to the safe action is reversible, information aggregation still fails when k is large, that is, when the ratio of votes required for the risky action is high.

Appendix 2.A Proofs

2.A.0.1 Proof of Proposition 2.1

Rewrite (2.6) as

$$s = p(kN-1,\hat{t})\lambda\left[z+\frac{g}{r}-\frac{s}{r}+(N-kN)(\frac{p(kN,\hat{t})g}{r}-\frac{s}{r})\right].$$

The right side is strictly decreasing in $\hat{t}$ whenever it is positive since $p(kN-1,\hat{t})$ is strictly decreasing in $\hat{t}$, as discussed in Section 2.2.3.

The right side grows without bound when $\hat{t}$ goes to 0 and N goes to infinity since

$$\lim_{\substack{N\to\infty\\ \hat{t}\to 0}} p(kN-1,\hat{t}) = \lim_{\substack{N\to\infty\\ \hat{t}\to 0}} p(kN,\hat{t}) = \rho_h,$$

and

$$\rho_h g > s.$$

The limit of the right side for any fixed N when $\hat{t}$ goes to infinity is non-positive since

$$p(kN-1,\hat{t}) < p(kN,\hat{t}) < p(\hat{t}|H),\ \forall N > 0 \text{ and } \forall\hat{t} > 0,$$

6. We can rule out the equilibria in which unsure voters vote for the risky action after the cut-off time characterized by the simple equilibria according to the discussion of requirement (iii) of Definition 2.1.

and

$$\lim_{\hat{t}\to\infty} p(\hat{t}|H) = 0.$$

Since the right side is continuous in $\hat{t}$, by the intermediate value theorem, we can find a unique solution of equation (2.6) when N is large enough, which ensures the existence and uniqueness of the simple equilibrium.

2.A.0.2 Proof of Proposition 2.2

We rewrite equation (2.6) by plugging in (2.5). The cut-off $\hat{t}_{k,N}$ is the solution of

$$\frac{q(kN-1,\hat{t})}{1-q(kN-1,\hat{t})}\cdot\frac{p(\hat{t}|H)g-s+p(\hat{t}|H)\lambda(\frac{g}{r}-\frac{s}{r})+(N-kN)p(\hat{t}|H)\lambda(\frac{p(\hat{t}|H)g}{r}-\frac{s}{r})}{s-p(\hat{t}|L)g-p(\hat{t}|L)\lambda(\frac{g}{r}-\frac{s}{r})-(N-kN)p(\hat{t}|L)\lambda(\frac{p(\hat{t}|L)g}{r}-\frac{s}{r})} = 1. \tag{2.A.1}$$

At each $t\in[0,\infty)$, there exists $\Lambda>0$ such that

$$\frac{1}{\Lambda} < \frac{\rho_H}{\rho_L}\frac{1-\rho_L(1-e^{-\lambda t})}{1-\rho_H(1-e^{-\lambda t})} < \Lambda,$$

$$\frac{1}{\Lambda} < \frac{\rho_L}{\rho_H}\frac{1-\rho_H(1-e^{-\lambda t})}{1-\rho_L(1-e^{-\lambda t})} < \Lambda.$$

Thus, for each $k\in(0,1]$ and each $N\in\mathbb{N}^+$,

$$\frac{1}{\Lambda}\frac{q_0}{1-q_0}\left\{\left(\frac{\rho_H}{\rho_L}\right)^k\left[\frac{1-\rho_H(1-e^{-\lambda t})}{1-\rho_L(1-e^{-\lambda t})}\right]^{1-k}\right\}^N < \frac{q(kN-1,t)}{1-q(kN-1,t)} < \Lambda\frac{q_0}{1-q_0}\left\{\left(\frac{\rho_H}{\rho_L}\right)^k\left[\frac{1-\rho_H(1-e^{-\lambda t})}{1-\rho_L(1-e^{-\lambda t})}\right]^{1-k}\right\}^N.$$

For each $k\in(0,1]$, we work with

$$\left\{\left(\frac{\rho_H}{\rho_L}\right)^k\left[\frac{1-\rho_H(1-e^{-\lambda t})}{1-\rho_L(1-e^{-\lambda t})}\right]^{1-k}\right\}^N.$$

Let $\hat{k}$ be that

$$\hat{k} = \frac{\ln\frac{1-\rho_L}{1-\rho_H}}{\ln\frac{\rho_H}{\rho_L}+\ln\frac{1-\rho_L}{1-\rho_H}}.$$

When $k<\hat{k}$, the equation

$$\left(\frac{\rho_H}{\rho_L}\right)^k\left[\frac{1-\rho_H(1-e^{-\lambda t})}{1-\rho_L(1-e^{-\lambda t})}\right]^{1-k} = 1$$

admits a unique solution. Denote it by $\hat{t}'_k$, which is strictly increasing in k. It follows that

$$\lim_{k\to 0}\hat{t}'_k = 0,$$
$$\lim_{k\to\hat{k}}\hat{t}'_k = \infty.$$

Set $\hat{t}'_k = \infty$ when $k \geq \hat{k}$. There exists $\bar{k} \in (0, \hat{k})$ such that $\hat{t}'_{\bar{k}} = \bar{t}$.

We now show that for $k < \bar{k}$, $\lim_{N\to\infty} \hat{t}_{k,N} = \hat{t}'_k$. If not, then there exist $\epsilon > 0$ and a sub-sequence of $\{1, 2, ...\}$ denoted by $\{n_1, n_2, ...\}$ such that

$$|\hat{t}_{k,n_i} - \hat{t}'_k| > \epsilon, \ \forall i \in \{1, 2, ...\}$$

Thus, for any $M > 0$, there exists $N(M)$ such that

$$\frac{q(kn_i - 1, t)}{1 - q(kn_i - 1, t)} \notin [\frac{1}{M}, M], \ \forall i > N(M).$$

Hence, the first term of (2.A.1) is extremely small or big when there are n_i voters and i is big. Since the second term of equation (2.A.1) is a bounded positive number when $k < \bar{k}$, the equation (2.A.1) is not valid — a contradiction.

Finally, we show that for $k \geq \bar{k}$, $\lim_{N\to\infty} \hat{t}_{k,N} = \bar{t}$. Suppose it is not true. Two cases might happen. One case is that there exist $\epsilon > 0$ and a sub-sequence of $\{1, 2, ...\}$ denoted by $\{n_1, n_2, ...\}$ such that

$$\hat{t}_{k,n_i} - \bar{t} > \epsilon, \ \forall i \in \{1, 2, ...\}$$

Thus, there exists $M < 0$ and $N(M) > 0$ such that the second term of the (2.A.1) is less than M when there are n_i voters with $i > N(M)$. Since the first term of (2.A.1) is positive, (2.A.1) is violated. The other case is that there exist $\epsilon > 0$ and a sub-sequence of $\{1, 2, ...\}$ denoted by $\{n_1, n_2, ...\}$ such that

$$\hat{t}_{k,n_i} - \bar{t} < -\epsilon, \ \forall i \in \{1, 2, ...\}.$$

Thus, for any $M > 0$, there exists $N(M)$ such that

$$\frac{q(kn_i - 1, t)}{1 - q(kn_i - 1, t)} > M, \ \forall i > N(M).$$

Hence, the first term of (2.A.1) is extremely big when there are n_i voters and i is big. As the second term of (2.A.1) is a positive number bounded away from 0 since $\hat{t}_{k,n_i} - \bar{t} < \epsilon$, (2.A.1) is violated.

2.A.0.3 Proof of Proposition 2.3

We define $\hat{k}$, $\hat{t}'_k$, $\bar{k}$ in the proof of Proposition 2.2.

From the proof of Proposition 2.2, we know that $\hat{t}_k = \hat{t}'_k$ for $k < \bar{k}$ and $\hat{t}_k = \bar{t}$ for $k \geq \bar{k}$. For each $k \in (0, 1]$, under the k—majority rule, by the law of large numbers, the ratios of sure winners at time $\hat{t}_{k,N}$ in both states converge in probability to fixed numbers when N goes to infinity, that is,

$$\frac{H(\hat{t}_{k,N})}{N} \xrightarrow{p} \rho_H(1 - e^{-\lambda \hat{t}_k}),$$

$$\frac{L(\hat{t}_{k,N})}{N} \xrightarrow{p} \rho_L(1-e^{-\lambda \hat{t}_k}).$$

Since $\hat{t}_k$ is bounded above by $\bar{t}$,

$$\rho_H(1-e^{-\lambda \hat{t}_k}) \le \rho_H(1-e^{-\lambda \bar{t}}) = \frac{\rho_H g - s}{g-s}.$$

Thus, information is not aggregated in state H for each $k \in [\frac{\rho_H g-s}{g-s}, 1]$. Note that $\bar{k} \le \frac{\rho_H g-s}{g-s} < \hat{k}$.

For each $k \in (0, \frac{\rho_H g-s}{g-s})$, we wish to prove that

$$\rho_H(1-e^{-\lambda \hat{t}_k}) > k > \rho_L(1-e^{-\lambda \hat{t}_k}).$$

Define

$$\begin{aligned} H(k) &= \rho_H(1-e^{-\lambda \hat{t}_k}), \\ L(k) &= \rho_L(1-e^{-\lambda \hat{t}_k}). \end{aligned}$$

It is direct to see that $H(k)$ and $L(k)$ are increasing and continuous in k when $k \le \bar{k}$. Both $H(k)$ and $L(k)$ stay constant when $k \in (\bar{k}, \frac{\rho_H g-s}{g-s}]$.

We can show that

$$\begin{aligned} \lim_{k\to 0} \frac{dH(k)}{dk} &= \frac{\rho_H}{\rho_H - \rho_L} \ln(\frac{\rho_H}{\rho_L}) > 1, \\ \lim_{k\to 0} \frac{dL(k)}{dk} &= \frac{\rho_L}{\rho_H - \rho_L} \ln(\frac{\rho_H}{\rho_L}) < 1. \end{aligned}$$

by choosing $x = \frac{\rho_H}{\rho_L}$ and using the fact that $1 - \frac{1}{x} < \ln(x) < x - 1$ for $x > 1$. Since $\lim_{k\to 0} H(k) = \lim_{k\to 0} L(k) = 0$, there exists $\tilde{k} > 0$ such that $H(k) > k > L(k)$ for $k \in (0, \tilde{k})$.

Now we prove $H(k) > k$ for $k \in (0, \bar{k}]$. It is enough to show that there does not exist $k \in (0, \bar{k})$ such that $H(k) = k$. If such k exists, then

$$\left(\frac{\rho_H}{\rho_L}\right)^k \left[\frac{1-\rho_H(1-e^{-\lambda \hat{t}_k})}{1-\rho_L(1-e^{-\lambda \hat{t}_k})}\right]^{1-k} = \left(\frac{\rho_H}{\rho_L}\right)^k \left[\frac{1-k}{1-\frac{\rho_L}{\rho_H}k}\right]^{1-k} > 1.$$

The inequality above is shown by setting $x = \frac{\rho_H}{\rho_L}$ and showing that $x^k(\frac{1-k}{1-\frac{k}{x}})^{1-k}$ is strictly increasing when $x > 1$ for each $k \in (0,1)$.

Next, there exists $\Lambda > 0$ such that for each $N > 0$,

$$\frac{1}{\Lambda} < \frac{\rho_H}{\rho_L} \frac{1-\rho_L(1-e^{-\lambda \hat{t}_{k,N}})}{1-\rho_H(1-e^{-\lambda \hat{t}_{k,N}})} < \Lambda,$$

$$\frac{1}{\Lambda} < \frac{\rho_L}{\rho_H} \frac{1-\rho_H(1-e^{-\lambda \hat{t}_{k,N}})}{1-\rho_L(1-e^{-\lambda \hat{t}_{k,N}})} < \Lambda.$$

Thus, for each $N \in \mathbb{N}^+$,

$$\frac{1}{\Lambda}\frac{q_0}{1-q_0}\left\{\left(\frac{\rho_H}{\rho_L}\right)^k\left[\frac{1-\rho_H(1-e^{-\lambda t})}{1-\rho_L(1-e^{-\lambda t})}\right]^{1-k}\right\}^N < \frac{q(kN-1,t)}{1-q(kN-1,t)} < \Lambda\frac{q_0}{1-q_0}\left\{\left(\frac{\rho_H}{\rho_L}\right)^k\left[\frac{1-\rho_H(1-e^{-\lambda t})}{1-\rho_L(1-e^{-\lambda t})}\right]^{1-k}\right\}^N.$$

Since

$$\lim_{N\to\infty} \hat{t}_{k,N} = \hat{t}_k,$$

$$\left(\frac{\rho_H}{\rho_L}\right)^k\left[\frac{1-\rho_H(1-e^{-\lambda t})}{1-\rho_L(1-e^{-\lambda t})}\right]^{1-k} > 1,$$

for each $M > 0$, there exists $N_M > 0$ such that for each $N > N_M$,

$$\frac{q(kN-1,\hat{t}_{k,N})}{1-q(kN-1,\hat{t}_{k,N})} > M.$$

Thus, the first term of equation (2.A.1) is extremely big when N is big. Since the second term of equation (2.A.1) is a bounded positive number when $k < \bar{k}$, equation (2.A.1) is violated. Therefore, there does not exist $k \in (0,\bar{k})$ such that $H(k) = k$. Hence, we have $H(k) > k$ for $k \in (0,\bar{k})$.

With the similar argument, we can prove $L(k) < k$ for $k \in (0,\bar{k})$.

Since $H(k)$ and $L(k)$ are continuous, we have

$$H(\bar{k}) \geq \bar{k} \geq L(\bar{k}).$$

Note that $H(\bar{k}) = \frac{\rho_H g-s}{g-s}$. Since $H(k)$ and $L(k)$ are constant when $k \geq \bar{k}$. It is directly to see that for $k \in (\bar{k}, \frac{\rho_H g-s}{g-s})$,

$$H(k) > k > L(k).$$

References

Austen-Smith, David, and Jeffrey S Banks. 1996. "Information aggregation, rationality, and the Condorcet jury theorem." *American Political Science Review* 90 (1): 34–45. [60]

Bolton, Patrick, and Christopher Harris. 1999. "Strategic experimentation." *Econometrica* 67 (2): 349–74. [60]

Duggan, John, and César Martinelli. 2001. "A Bayesian model of voting in juries." *Games and Economic Behavior* 37 (2): 259–94. [59, 67]

Feddersen, Timothy, and Wolfgang Pesendorfer. 1997. "Voting behavior and information aggregation in elections with private information." *Econometrica* 65 (5): 1029–58. [59, 60, 67]

Feddersen, Timothy, and Wolfgang Pesendorfer. 1998. "Convicting the innocent: The inferiority of unanimous jury verdicts under strategic voting." *American Political Science Review* 92 (1): 23–35. [60]

Fernandez, Raquel, and Dani Rodrik. 1991. "Resistance to reform: Status quo bias in the presence of individual-specific uncertainty." *American Economic Review* 81 (5): 1146–55. [60, 68]

Halac, Marina, Navin Kartik, and Qingmin Liu. 2017. "Contests for experimentation." *Journal of Political Economy* 125 (5): 1523–69. [60]

Keller, Godfrey, Sven Rady, and Martin Cripps. 2005. "Strategic experimentation with exponential bandits." *Econometrica* 73 (1): 39–68. [60]

Moldovanu, Benny, and Frank Rosar. 2021. "Brexit: A comparison of dynamic voting games with irreversible options." *Games and Economic Behavior* 130 (5): 85–108. [60]

Murto, Pauli, and Juuso Välimäki. 2011. "Learning and information aggregation in an exit game." *Review of Economic Studies* 78 (4): 1426–61. [60]

Strulovici, Bruno. 2010. "Learning while voting: Determinants of collective experimentation." *Econometrica* 78 (3): 933–71. [58, 59, 69]

Chapter 3

Fishing for Approval

Joint with Mehmet Ekmekci and Stephan Lauermann

3.1 Introduction

One success can make up for a long record of failure. For example, a lobby is searching for an expert to support its position publicly. Once it finds an advocate, none of those past attempts matter anymore. Other examples include an international student who takes the TOEFL repeatedly until he receives a satisfying grade, or a family who visit charities sequentially until they convince one of them that they are in need. Such “fishing for approval" happens generally when someone needs a critical assent and engages in a costly search to solicit one from decentralized approval agencies. Approval agencies often have common interests, in that they only want the qualified person to receive approval. However, fishing for approval seems to be a substantial problem for them. No matter how many agencies give rejections under contemplative reasoning, an unqualified person can still turn the tide by fishing for one fortuitous approval.

Here, we analyze such case of fishing for approval within the framework of two-option elections. An organizer samples voters sequentially to vote on two policies, a and b. He incurs a cost for each voter sampled. If one voter votes for a, then the game ends and a is implemented. If the organizer stops sampling or he exhausts the pool of voters, then b is implemented. Voters have common interests and prefer a in state α and b in state β. However, none of them knows the realized state. They only obtain noisy information as private signals. The organizer is informed about the realized state and prefers a regardless of it. He is fishing for approval for a.

This paper studies a sequential voting model under unanimity rule. Feddersen and Pesendorfer (1998) and Duggan and Martinelli (2001) show that unanimity rule produces asymptotic inefficiency in simultaneous common-value elections with an exogenous number of voters. Even in large elections, unanimity rule generally fails to aggregate dispersed information among voters (unless there exists an unboundedly informative signal). Dekel and Piccione (2000) show that the whole set of equilibria is the same in all sequential voting structures under the unanimity rule. Hence, sequential voting structures do not facilitate the aggregation of private information.

Our paper raises additional concern: by fishing for approval, the biased organizer might manipulate the voting outcome in his favor. However, we show that, somewhat paradoxically, rather than hurting, the organizer's ability to fish for approval helps voters. Since the organizer bases his sampling behavior on his knowledge of the state, voters can "manipulate" him to reveal his private information by generating state-dependent approval probability. The endogeneity of the sequential voting structure is the key difference between this paper and Dekel and Piccione (2000). Our main result Proposition (3.2) shows that voters can reach their first-best outcome as the number of voters goes to infinity.

Information aggregation is a consequence of the interplay between the incentives of the organizer to be willing to sample everyone in state α and the increasingly negative inference the voters draw from conditioning on other voters' disapproval. A voter's approval probability must be high enough to make up for the organizer's sampling cost. Hence, disapproval by one voter is an informative signal in favor of state β. A voter votes according to her inference from being pivotal, that is, on disapproval from all other voters. As the number of voters grows, if the organizer sampled in state β as well, then a voter would put a sufficiently large weight on state β when conditioning on being pivotal, and would approve with a vanishing probability. This, however, is inconsistent with the organizer's sampling incentives. Hence, in equilibrium, the organizer samples in state α with probability 1 and only with a vanishing probability in state β.

This paper proceeds as follows: Section 3.2 describes the basic model in which voters do not observe their order in the sampling sequence. Section 3.3 characterizes the equilibrium and presents our main result in large elections. Section 3.4 analyzes the observable-order case. Section 3.5 concludes.

3.2 Model

3.2.1 Model Setting

An *organizer* samples among $M > 1$ *voters* sequentially. He incurs a sampling cost $c \in (0, 1)$ for each voter sampled. Sampling is random without replacement. Voters do not observe their order in the sampling sequence.

After a voter is sampled, she votes on two policies, $\{a, b\}$. If she votes for a, that is, she approves a, then the game ends and a is implemented. If she votes for b, that is, she rejects a, then the organizer decides whether to continue sampling. Policy b is implemented if there is no approval for a, which happens if (i) the organizer stops initially without sampling anyone, (ii) he stops sampling after one voter rejects a, or (iii) he has sampled all M voters and none of them approves a.

Voters' payoffs depend on the realized state $\omega \in \{\alpha, \beta\}$. They share the following utility function:

$$u(a,\alpha) = u(b,\beta) = 1,$$
$$u(a,\beta) = u(b,\alpha) = 0,$$

where $u(x,\omega)$ denotes the utility if the policy x is chosen in state ω.

Voters are uncertain about the realized state ω. They share a common prior belief $\pi \in (0,1)$ that the state is α. Each voter $i \in \{1,\ldots,M\}$ receives a private signal $s_i \in [0,1]$. Conditional on the state, signals are independent across voters. In state $\omega \in \{\alpha,\beta\}$, signals are distributed according to a cumulative distribution function $F(s|w)$ with a continuous density function $f(s|w)$.

We assume the strict ***Monotone Likelihood Ratio Property (MLRP)***:

$$\frac{f(s|\alpha)}{f(s|\beta)} \quad \text{is strictly decreasing in s.} \tag{3.1}$$

Assumption (3.1) implies that higher signals are stronger indicators of state β. It also implies that $F(s|\alpha) > F(s|\beta)$ for each $s \in (0,1)$.

We further assume that[1]

$$\lim_{s\to 1} \frac{\pi}{1-\pi}\frac{f(s|\alpha)}{f(s|\beta)} < 1. \tag{3.2}$$

The organizer observes the realized state ω and prefers a to be implemented in both states. If in total n voters are sampled, then his payoff is:

$$u_o(a) = 1 - cn,$$
$$u_o(b) = -cn.$$

1. Duggan and Martinelli (2001) specify assumption (3.2) to ensure the existence of a non-trivial equilibrium in the simultaneous common-value voting game. We can replace assumption (3.2) by assuming $\lim_{s\to 1} \frac{\pi}{1-\pi}\left(\frac{f(s|\alpha)}{f(s|\beta)}\right)^M < 1$ without changing any result.

3.2.2 Strategy

We examine symmetric *Bayesian Nash equilibrium*, in which all voters use the same cut-off strategy[2] $\hat{s} \in [0,1]$. Each voter $i \in \{1,...,M\}$ approves a if $s_i < \hat{s}$ and rejects a if $s_i \geq \hat{s}$. In state $w \in \{\alpha, \beta\}$, one voter's approval probability is $F(\hat{s}|\omega)$.

The strategy of the organizer is a function

$$p \; : \; \{0,...,M\} \times \{\alpha,\beta\} \; \rightarrow \; [0,1],$$

such that

$$\sum_{n=0}^{M} p(n,\omega) \; = \; 1, \; \forall \omega \; \in \; \{\alpha,\beta\}.$$

With probability $p(n,\omega)$, the organizer samples *up to n* voters in state $\omega \in \{\alpha,\beta\}$, he stops if (i) one voter approves a, or (ii) he has sampled n voters and no one approves a.

3.2.3 Preliminary Analysis

<u>Inference of Voters</u>

Each voter sampled is *pivotal* if (i) the organizer stops sampling after she rejects a, or (ii) all voters sampled after her reject a. When deciding how to vote, it is optimal to condition on the event that she is pivotal since her vote cannot affect the outcome in any other event.

Assume that voters use a cut-off strategy $\hat{s}$ and the organizer uses a pure strategy p with $p(n_\alpha,\alpha) = 1$ and $p(n_\beta,\beta) = 1$ for some $n_\alpha, n_\beta > 0$. Consider a voter who is sampled and pivotal in state $\omega \in \{\alpha,\beta\}$. (i) The voters sampled before her must have rejected a, otherwise she is not sampled. (ii) Since she is pivotal, if she rejects a, then all voters sampled after her reject a. Thus, if she rejects a, then in state ω the organizer samples in total n_ω voters and the remaining $n_\omega - 1$ voters reject a. The likelihood of being sampled and pivotal in state ω is

$$\frac{n_\omega}{M}(1 - F(\hat{s}|\omega))^{n_\omega - 1}.$$

For each voter, when the strategy of the organizer is p and the strategies of other voters are $\hat{s}$, the posterior likelihood ratio that the state is α, conditional on (i) receiving a signal s, (ii) being sampled, and (iii) being pivotal, is denoted by $\Phi(s, \mathrm{spl}, \mathrm{piv}; p, \hat{s}, M)$. We suppress arguments $p, \hat{s}, M$ and denote it by

$$\Phi(s,spl,piv) \; = \; \underbrace{\frac{\pi}{1-\pi}}_{\text{prior}} \; \underbrace{\frac{f(s|\alpha)}{f(s|\beta)}}_{\text{signal}} \; \underbrace{\frac{\sum_{m=1}^{M} p(m,\alpha) \; \frac{m}{M} \; (1-F(\hat{s}|\alpha))^{m-1}}{\sum_{n=1}^{M} p(n,\beta) \; \frac{n}{M} \; (1-F(\hat{s}|\beta))^{n-1}}}_{\text{sampled and pivotal under distribution } p}. \tag{3.3}$$

2. In Appendix 3.A.3 of the supplement material, we show that focusing on the cut-off strategy is without loss of generality from the *MLRP* assumption.

If the denominator is 0 but the numerator is not, then we set $\Phi(s, spl, piv) = \infty$. If both the denominator and the numerator are 0, then we set $\Phi(s, spl, piv) = 1$. Each voter approves a if $\Phi(s, spl, piv) > 1$ and rejects a if $\Phi(s, spl, piv) < 1$. She is indifferent between a and b if $\Phi(s, spl, piv) = 1$.

Organizer's Sampling Problem

The organizer's sampling problem is stationary. It only depends on the cost c and the approval probability $F(\hat{s}|\omega)$ in state $\omega \in \{\alpha, \beta\}$. The strategy p is optimal if

$$\begin{cases} p(0,\omega) = 1 & \text{when } F(\hat{s}|\omega) < c, \\ p(M,\omega) = 1 & \text{when } F(\hat{s}|\omega) > c, \\ (p(0,\omega), \ldots., p(M,\omega)) \in \Delta\{[0,1]^{M+1}\} & \text{when } F(\hat{s}|\omega) = c. \end{cases} \tag{3.4}$$

In the first case, the organizer stops initially. In the second case, the organizer keeps sampling voters. In the third case, the organizer is indifferent between all stopping time.

3.3 Equilibrium Analysis

3.3.1 Characterization

An equilibrium is *responsive* if (i) the organizer does active sampling in both states,

$$p(0,\alpha) < 1 \text{ and } p(0,\beta) < 1,$$

and (ii) voters do not always approve a, that is, $\hat{s} < 1$.

The first condition rules out the trivial equilibrium in which the organizer stops initially in both states. In such equilibrium, each voter uses the cut-off strategy $\hat{s}$ such that $F(\hat{s}|\alpha) \leq c$ when being sampled. Policy b is always implemented. In addition, there does not exist any equilibrium[3] where the organizer only samples voters in one state. If the organizer only samples voters in state α, then all voters always approve a, but this would also induce the organizer to sample voters in state β. If the organizer only sample voters in state β, then all voters always reject a, but this would induce the organizer to stop initially in both states.

The second condition rules out the trivial equilibrium in which all voters always approve a with probability 1 when being sampled. In such equilibrium, the organizer keeps sampling voters in both states since $c < 1$. Since all voters always approve a and there are more than one voters, none of them is pivotal. Policy a is always implemented.

3. When referring to equilibrium without further qualification, we mean symmetric equilibria.

In a responsive equilibrium, since the organizer does active sampling in both states, from (3.4),

$$F(\hat{s}|\alpha) \; > \; F(\hat{s}|\beta) \; \geq \; c. \tag{3.5}$$

The strict inequality comes from the *MLRP* assumption. Since $F(\hat{s}|\alpha) > c$, the organizer keeps sampling voters in state α,

$$p(M, \alpha) \; = \; 1. \tag{3.6}$$

In state β,

$$p(M, \beta) \; = \; 1 \qquad \text{when } F(\hat{s}|\beta) \; > \; c, \tag{3.7}$$

$$(p(0, \beta), ..., p(M, \beta)) \; \in \; \Delta\{[0, 1]^{M+1}\} \qquad \text{when } F(\hat{s}|\beta) \; = \; c. \tag{3.8}$$

Denote the vector $(p(0, \beta), ..., p(M, \beta))$ by p_β.

Let $\hat{s}_c$ such that $F(\hat{s}_c|\beta) = c$. From (3.5) and the condition that $\hat{s} < 1$,

$$\hat{s} \; \in \; [\hat{s}_c, 1). \tag{3.9}$$

Voters are indifferent between a and b when receiving the cut-off signal $\hat{s}$. Plug in (3.6) to (3.3), the cut-off $\hat{s}$ is determined by

$$\Phi(\hat{s}, spl, piv) \; = \; \frac{\pi}{1-\pi} \frac{f(\hat{s}|\alpha)}{f(\hat{s}|\beta)} \frac{(1-F(\hat{s}|\alpha))^{M-1}}{\sum_{n=1}^{M} p(n, \beta) \; \frac{n}{M} \; (1-F(\hat{s}|\beta))^{n-1}} \; = \; 1. \tag{3.10}$$

A responsive equilibrium is characterized by a pair $(p_\beta, \hat{s})$ satisfying (3.7), (3.8), (3.9) and (3.10).

Proposition 3.1. *A responsive equilibrium exists for each* $c \in (0, 1)$ *and* $M > 1$.

In the Appendix 3.A.1, we construct one responsive equilibrium in which the organizer mixes: he either stops initially or keeps sampling voters in state β, that is,

$$p(0, \beta) + p(M, \beta) \; = \; 1. \tag{3.11}$$

We denote this equilibrium as the *simple equilibrium*. We show that a unique simple equilibrium exists for each $c \in (0, 1)$ and $M > 1$.

3.3.2 Large Elections

For each cost $c \in (0,1)$, consider a sequence of responsive equilibria corresponding to M. The following proposition shows that every sequence of responsive equilibria leads to voters' first-best outcome as $M \to \infty$.

Proposition 3.2. *For each $c \in (0,1)$, the voter's expected payoff converges to 1 in every sequence of responsive equilibria as $M \to \infty$.*

We prove Proposition 3.2 by two lemmas. Lemma 3.1 characterizes voters' equilibrium strategy when M is large.

Lemma 3.1. *For each $c \in (0,1)$, there exists $M_1(c)$ such that for each $M > M_1(c)$, voters use the cut-off strategy $\hat{s}_c$ in every responsive equilibrium.*

Proof. Fix the cost c. Consider a responsive equilibrium in which voters do not use the cut-off strategy $\hat{s}_c$. If $\hat{s} \neq \hat{s}_c$, then $\hat{s} > \hat{s}_c$ from (3.9). Using (3.5), we have $p(M,\beta) = 1$. Hence, we can rewrite (3.10) as

$$L(\hat{s};M) \;\equiv\; \frac{\pi}{1-\pi}\,\frac{f(\hat{s}|\alpha)}{f(\hat{s}|\beta)}\,\frac{(1-F(\hat{s}|\alpha))^{M-1}}{(1-F(\hat{s}|\beta))^{M-1}}$$

If such equilibrium exists, then $L(\hat{s};M) = 1$ for some $\hat{s} > \hat{s}_c$.

By the *MLRP* assumption, the function $L(s;M)$ is strictly decreasing in s. Hence,

$$L(\hat{s};M) \;<\; L(\hat{s}_c;M).$$

Since $(1-F(\hat{s}_c|\alpha)) < (1-F(\hat{s}_c|\beta))$,

$$\lim_{M\to\infty} L(\hat{s}_c;M) \;=\; 0.$$

Thus, we can pick $M_1(c)$ such that for each $M > M_1(c)$, the likelihood ratio $L(\hat{s};M)$ is strictly less than 1 for each $\hat{s} \in [\hat{s}_c, 1)$[4]. Therefore, for each $M > M_1(c)$, voters use the cut-off strategy $\hat{s}_c$ in every responsive equilibrium. □

Lemma 3.2 characterizes the organizer's equilibrium strategy. The organizer hardly samples any voters in state β when M is large.

Lemma 3.2. *For each $c \in (0,1)$, $p(0,\beta)$ converges to 1 in every sequence of responsive equilibria as $M \to \infty$.*

4. There exists $\delta > 0$ such that $\lim_{\hat{s}\to 1} L(\hat{s};M) < 1-\delta$ for each $M > 1$ according to assumption (3.2) and L'Hôpital's rule.

Proof. Fix the cost c. For each $M > M_1(c)$, by Lemma 3.1, voters use the cut-off strategy $\hat{s}(c)$ in every responsive equilibrium. Rewrite (3.10) as

$$\Phi(\hat{s}_c, spl, piv) = \frac{\pi}{1-\pi}\frac{f(\hat{s}_c|\alpha)}{f(\hat{s}_c|\beta)}\frac{(1-F(\hat{s}_c|\alpha))^{M-1}}{\sum_{n=1}^{M} p(n,\beta)\,\frac{n}{M}\,(1-F(\hat{s}_c|\beta))^{n-1}} = 1. \quad (3.12)$$

Consider the function $g(x) = xa^x$ with $a \in (0,1)$. Note that (i) $g(x)$ is decreasing in x when $x > -\frac{1}{\log a}$, and (ii) $\lim_{x\to\infty} g(x) = 0$. Thus, there exists $\bar{x}(a) > 0$ such that for each $\hat{x} > \bar{x}(a)$,

$$\hat{x} = \arg\min_{x\in[1,\hat{x}]} g(x).$$

Hence, there exists $M_2(c)$ such that for each $M > M_2(c)$,

$$(1-F(\hat{s}_c|\beta))^{M-1} < \frac{n}{M}(1-F(\hat{s}_c|\beta))^{n-1}, \ \forall n \in \{1,\ldots,M-1\}. \quad (3.13)$$

For each $M > \max\{M_1(c), M_2(c)\}$, using (3.12) and (3.13)

$$1 = \Phi(\hat{s}_c, spl, piv) \leq \frac{\pi}{1-\pi}\frac{f(\hat{s}_c|\alpha)}{f(\hat{s}_c|\beta)}\frac{(1-F(\hat{s}_c|\alpha))^{M-1}}{(1-p(0,\beta))(1-F(\hat{s}_c|\beta))^{M-1}}.$$

Therefore,

$$p(0,\beta) \geq 1 - \frac{\pi}{1-\pi}\frac{f(\hat{s}_c|\alpha)}{f(\hat{s}_c|\beta)}\frac{(1-F(\hat{s}_c|\alpha))^{M-1}}{(1-F(\hat{s}_c|\beta))^{M-1}}. \quad (3.14)$$

The right side converges to 1 when $M \to \infty$ since $(1-F(\hat{s}_c|\alpha)) < (1-F(\hat{s}_c|\beta))$. □

In every sequence of responsive equilibria, in state α, when $M > M_1(c)$, the probability of implementing a is $1-((1-F(\hat{s}_c|\alpha)))^M$, which converges to 1. In state β, the probability of implementing b converges to 1 since $p(0,\beta)$ converges to 1. Thus, every sequence of responsive equilibria leads to voters' first-best outcome as $M \to \infty$.

Proposition 3.2 seems counter-intuitive since the biased organizer should manipulate the election outcome. However, since he bases his sampling behavior on his knowledge of the state, voters can "manipulate" him to reveal his private information by generating state-dependent approval probability. The organizer is indifferent in state β when voters use the cut-off strategy $\hat{s}_c$. Note that $\hat{s}(c)$ is determined by the cost c, independent of the number of voters M. Since $(1-F(\hat{s}_c|\alpha))^{M-1}$ goes to 0 much faster than $\frac{n}{M}(1-F(\hat{s}_c|\beta))^{M-1}$ for each $n \leq M$ when $M \to \infty$, from (3.12), the organizer has to choose $p(0,\beta)$ near 1 to keep the likelihood ratio away from 0.

Remark 3.1. Every sequence of responsive equilibria converges to the same strategy profile in which (i) voters use the cut-off strategy $\hat{s}_c$, and (ii) the organizer keeps sampling voters in state α and stops initially in state β. This strategy profile is an equilibrium leading to voters' first-best outcome when there are infinitely many voters. In addition, this strategy profile maximizes the expected payoff of the organizer among all the equilibria leading to voters' first-best outcome since it maximizes the approval probability in state α.

Remark 3.2. The inequality (3.14) is an equality in the simple equilibrium. Hence, when M is large, the probability of implementing b in state β is lowest in the simple equilibrium among all responsive equilibria. The simple equilibrium is the worst responsive equilibrium for voters when M is large. In Appendix 3.A.4 of the supplement material, we show that the simple equilibrium is the worst responsive equilibrium for all $M > 1$.

Remark 3.3. We can construct an asymmetric equilibrium in which voters gain their first-best outcome when M is above some threshold by (i) letting one voter approve a with probability 1 to impede others from being pivotal, and (ii) letting other voters use the same cut-off strategy $\hat{s} \in (0, 1)$ to make the organizer keep sampling voters in state α and stop initially in state β. However, this equilibrium requires a strong collaboration between voters and is not robust even when there is a tiny uncertainty about the total number of voters. For example, assume that the game ends with probability $\epsilon > 0$ after each voter votes. No matter how small ϵ is, if the organizer only samples voters in state α, for each voter, she is pivotal with positive probability and always approves a, which would induce the organizer to sample voters in state β. Additionally, if we assume the number of voters follows a Poisson distribution with mean M, such asymmetric equilibrium does not exist while Proposition 3.2 remains valid.

3.4 Observable Order

In the basic model, voters do not observe their order in the sampling sequence. Now, we analyze the model in which voters observe their order. When someone is searching for approval from decentralized agencies, they might know how many rejections he has received if the privacy requirement is weak.

3.4.1 Symmetric Equilibrium

We first focus on symmetric *Bayesian Nash equilibrium* in which voters use the same cut-off strategy $\hat{s} \in [0, 1]$ regardless of their order. The strategy of the organizer is a function p defined in Section 3.2.2. Compared to the basic model with unobservable order, the organizer faces the same sampling problem, which is discussed in Section 3.2.3. The only difference is the inference of voters.

Consider the n^{th} voter in the sequence with $n \in \{1, ..., M\}$. Given that the organizer uses the strategy p and other voters use the cut-off strategy $\hat{s}$, if she is sampled in state $\omega \in \{\alpha, \beta\}$, then (i) voters before her must have rejected a, and (ii) the organizer does not stop before sampling her. Hence, the likelihood of being sampled in state ω is

$$\underbrace{(1-F(\hat{s}|\omega))^{n-1}}_{\text{(i)}}\underbrace{(1-\sum_{i=0}^{n-1}p(i,\omega))}_{\text{(ii)}}.$$

If she is pivotal, then voters sampled after her must reject a. The likelihood of being pivotal in state ω is

$$\sum_{j=n}^{M}\underbrace{\frac{p(j,\omega)}{1-\sum_{i=0}^{n-1}p(i,\omega)}}_{\text{distribution}}\underbrace{(1-F(\hat{s}|\omega))^{j-n}}_{\text{all reject } a}.$$

Hence, for the n^{th} voter, the likelihood of being sampled and pivotal in state $\omega\in\{\alpha,\beta\}$ is

$$\sum_{j=n}^{M}p(j,\omega)(1-F(\hat{s}|\omega))^{j-1}.$$

For the n^{th} voter, when the strategy of the organizer is p and the strategies of other voters are $\hat{s}$, the posterior likelihood ratio that the state is α, conditional on (i) receiving a signal s, (ii) being sampled, and (iii) being pivotal, is

$$\Phi_n(s,spl,piv) = \underbrace{\frac{\pi}{1-\pi}}_{\text{prior}}\underbrace{\frac{f(s|\alpha)}{f(s|\beta)}}_{\text{signal}}\underbrace{\frac{\sum_{j=n}^{M}p(j,\alpha)(1-F(\hat{s}|\alpha))^{j-1}}{\sum_{j=n}^{M}p(j,\beta)(1-F(\hat{s}|\beta))^{j-1}}}_{\text{sampled and pivotal}}. \tag{3.15}$$

We now characterize the responsive equilibrium[5] defined in Section 3.3.1. In a responsive equilibrium, since the organizer faces the same sampling problem, from the discussion in Section 3.3.1, he keeps sampling voters in state α with $p(M,\alpha)=1$. In state β, his equilibrium strategy is still characterized by (3.7) and (3.8). For voters, since they are indifferent between a and b when receiving the cut-off signal $\hat{s}$. Plug in $p(M,\alpha)=1$ to (3.15), the cut-off $\hat{s}$ is determined by

$$\Phi_n(\hat{s},spl,piv) = \frac{\pi}{1-\pi}\frac{f(\hat{s}|\alpha)}{f(\hat{s}|\beta)}\frac{(1-F(\hat{s}|\alpha))^{M-1}}{\sum_{j=n}^{M}p(j,\beta)(1-F(\hat{s}|\beta))^{j-1}} = 1,\ \forall n\in\{1,\dots,M\}. \tag{3.16}$$

A responsive equilibrium is characterized by (3.7), (3.8), (3.9) and (3.16). Compared to the characterization of the responsive equilibrium in the basic model, we only replace (3.10) by (3.16).

Proposition 3.3. *When the order is observable, a unique responsive equilibrium exists for each $c\in(0,1)$ and $M>1$, which is the simple equilibrium in the basic model.*

5. For equilibria which are not responsive, the two types of trivial equilibria mentioned in Section Section 3.3.1 are still valid equilibria when the order is observable.

Proof. From (3.16),

$$p(n,\beta) \propto \frac{1}{\Phi_n(\hat{s}, spl, piv)} - \frac{1}{\Phi_{n+1}(\hat{s}, spl, piv)} = 0, \ \forall n \in \{1, ..., M-1\}.$$

Hence, in every responsive equilibrium, we must have $p(n,\beta) = 0$ for each $n \in \{1, ..., M-1\}$, that is, we must have $p(0,\beta) + p(M,\beta) = 1$. Plug in this to (3.16),

$$\frac{\pi}{1-\pi} \frac{f(\hat{s}|\alpha)}{f(\hat{s}|\beta)} \frac{(1-F(\hat{s}|\alpha))^{M-1}}{(1-p(0,\beta))(1-F(\hat{s}|\beta))^{M-1}} = 1. \tag{3.17}$$

In the proof of Proposition 3.1, we show that there exits a unique strategy profile satisfying (3.7), (3.8), (3.9) and (3.17), which is the simple equilibrium in the basic model. Hence, the simple equilibrium is the unique responsive equilibrium when the order is observable. □

Remark 3.4. From Remark 3.2, in the basic model with unobservable order, the simple equilibrium is the worst responsive equilibrium. Hence, voters are worse off when the order is observable. Knowing the order does not help one voter know more about other voters' private information since she already makes decisions based on being sampled. Instead, voters are worse off since the available strategies of the organizer are restricted. It is harder for the voters to "manipulate" the organizer.

From Proposition 3.2 and Proposition 3.3, the sequence of responsive equilibria leads to voters' first-best outcome as $M \to \infty$,

Corollary 3.1. *When the order is observable, for each $c \in (0,1)$, the voter's expected payoff converges to 1 in the sequence of responsive equilibria as $M \to \infty$.*

3.4.2 Asymmetric Equilibrium

We now analyze the asymmetric equilibrium where voters with different orders might choose different strategies. The strategy profile of voters is characterized by a sequence of cut-offs $\{\hat{s}_1, ..., \hat{s}_M\}$ where the n^{th} voter uses the cut-off strategy $\hat{s}_n$ for each $n \in \{1, ..., M\}$. The strategy of the organizer is still characterized by function p. We focus on *Bayesian Nash equilibrium*.

We can construct an asymmetric equilibrium in which voters gain their first-best outcome when M is above some threshold by (i) letting the last voter approve a with probability 1 to impede others from being pivotal, and (ii) letting other $M-1$ voters use the same cut-off strategy $\hat{s} \in (0,1)$ to make the organizer keep sampling voters in state α and stop initially in state β. However, these equilibria are not robust even when there is a tiny uncertainty about the total number of voters, as discussed in Remark 3.3.

An equilibrium is *non-trivial* if (i) each voter does not always approve a and (ii) the organizer keep sampling voters in state α, that is,

$$\begin{aligned} \hat{s}_n &< 1, \qquad \forall n \in \{1, ..., M\}, \\ p(M, \alpha) &= 1. \end{aligned}$$

The second condition is based on D1 refinement. The organizer benefits more by sampling one more voter in state α for each $\hat{s} \in [0, 1]$ this voter chooses. Hence, for a voter who is never sampled in state α and β on the equilibrium path, if she is sampled, she believes that the state must be α and approves a.

The following proposition shows that every sequence of non-trivial equilibria leads to voters' first-best outcome as $M \to \infty$.

Proposition 3.4. *For each $c \in (0, 1)$, the voter's expected payoff converges to 1 in every sequence of non-trivial equilibria as $M \to \infty$.*

The proof is similar to the one of Proposition 3.2. Consider a sequence of non-trivial equilibria. First, when M is large, for each voter, either (i) the probability that the organizer stops sampling in state β before reaching her[6] approaches 1, or (ii) the probability that she approves a approaches or equals 0[7], since the posterior belief that the state is β conditional on rejections from all other voters converges to 1 as $M \to \infty$. Hence, the probability that the organizer receives the approval in state β from each voter[8] converges to 0 as $M \to \infty$. Second, we can find $T_c \in \mathbb{N}^+$ independent to M such that one of the first T_c voters must approve a with probability bounded away from 0 in both states. Otherwise, the organizer stops initially in both states. Third, when $M \to \infty$, the probability that the organizer stops sampling before reaching the $(T_c + 1)^{th}$ voter in state β must converges to 1. Otherwise, we can construct a subsequence of non-trivial equilibria where the organizer samples all the first T_c voters with probability bounded away from 0 when $M \to \infty$. According to the first argument, all the the first T_c voters must approve a with probability converging to 0, which contradicts the second argument. Therefore, the probability of implementing b in state β converges to 1.

The observable-order case relates to the social learning literature (Banerjee (1992), Bikhchandani, Hirshleifer, and Welch (1992)). This literature analyzes a sequential decision model in which each decision-maker looks at the decisions made by previous decision-makers in taking her own decision. This setting can lead rapidly to an inefficient "herd-cascade" in which subsequent decision-makers optimally ignore their private information and imitate earlier decision-makers. However, the "herd-cascade" does not happen in our model since voters learn

6. which is $\sum_{j=0}^{i-1} p(j, \beta)$ for the i^{th} voter.
7. It depends on whether $\lim_{s \to 0} \frac{f(s|\alpha)}{f(s|\beta)} = \infty$.
8. which is $\left[1 - \sum_{j=0}^{i-1} p(j, \beta)\right] \cdot \prod_{j=0}^{i-1} \left[1 - F(\hat{s}_j|\beta)\right] \cdot F(\hat{s}_i|\beta)$ for the i^{th} voter.

not only from the past, that is, conditional on being sampled, but also from the future, that is, conditional on being pivotal[9].

3.5 Conclusion

There is a concern that a biased agent might fish for approval if one assent undoes all past rejections. In this paper, we study a sequential voting model in which a biased organizer engages in a costly search to solicit one vote for his preferred policy. We show that the ability to fish for approval actually helps voters. By basing his sampling behavior on his knowledge of the state, the organizer ends up injecting additional information, often leading to the voters' first-best, full-information equivalent outcome.

This paper suggests several promising directions for future research. (1) We expect our main result can be extended to the interdependent-value case and other extreme rules under which a fixed number of approval are required. (2) It would be interesting to consider the costly voting here, following Krishna and Morgan (2011) and Krishna and Morgan (2015). The cost of voting corresponds to the expense of providing support when someone is searching for help among decentralized agencies. We could also consider the cost of information acquisition, following Martinelli (2006). (3) We analyze the case where the organizer is informed about the realized state. It is natural to study whether information can be aggregated if the organizer is only partially informed about the realized state.

Appendix 3.A Proofs

The appendices proceed as follows:

(1) In Appendix 3.A.1, we prove Proposition 3.1 and characterize the simple equilibrium.
(2) In Appendix 3.A.2, we prove Proposition 3.4 for the case where the voters can observe the order.
(3) In Appendix 3.A.3, we characterize the set of symmetric Bayesian Nash equilibrium and show that we can focus on the cut-off strategy without loss of generality.
(4) In Appendix 3.A.4, we characterize the set of responsive equilibrium and show that the simple equilibrium is the worst responsive equilibrium for the voters.

9. Ali and Kartik (2012) study sequential voting and identifies voting equilibria with herding. Our model's main difference from them is the endogeneity of the sampling process, hence the number of voters.

3.A.1 Proof of Proposition 3.1

We now show the existence and the uniqueness of the simple equilibrium for each $c \in (0,1)$ and $M > 1$.

If a strategy of the organizer p_{sim} is a part of a simple equilibrium, then it must satisfy (3.11). Plug in (3.11) to (3.10),

$$\Phi_{sim}(\hat{s}, spl, piv) = \frac{\pi}{1-\pi} \frac{f(\hat{s}|\alpha)}{f(\hat{s}|\beta)} \frac{(1-F(\hat{s}|\alpha))^{M-1}}{(1-p(0,\beta))\ (1-F(\hat{s}|\beta))^{M-1}}. \tag{3.A.1}$$

A simple equilibrium is characterized by (3.7), (3.8), (3.9) and (3.A.1).

Define

$$L(s) \equiv \frac{\pi}{1-\pi} \frac{f(s|\alpha)}{f(s|\beta)} \left(\frac{1-F(s|\alpha)}{1-F(s|\beta)}\right)^{M-1}.$$

It follows that

$$\Phi_{sim}(s, spl, piv) = L(s) \frac{1}{1-p_{sim}(0,\beta)}. \tag{3.A.2}$$

The function $L(s)$ is continuous in s since the density functions are continuous. It is strictly decreasing in s due to the *MLRP* assumption. By assumption (3.2),

$$\lim_{s\to 1} L(s) = \lim_{s\to 1} \frac{\pi}{1-\pi} \left(\frac{f(s|\alpha)}{f(s|\beta)}\right)^{M} < 1.$$

We also know that $\lim_{s\to 0} L(s) = \lim_{s\to 0} \frac{\pi}{1-\pi} \frac{f(s|\alpha)}{f(s|\beta)}$. If $\lim_{s\to 0} \frac{\pi}{1-\pi} \frac{f(s|\alpha)}{f(s|\beta)} > 1$, then for each $M \in \mathbb{N}^+$, there exists a unique $s \in (0,1)$ such that $L(s) = 1$, denoted by $s^*(M)$. If $\lim_{s\to 0} \frac{\pi}{1-\pi} \frac{f(s|\alpha)}{f(s|\beta)} \leq 1$, then let $s^*(M)$ equal 0.

Consider a pair of $c \in (0,1)$ and $M > 1$.

First, consider the case where $c \leq F(s^*(M)|\beta)$, that is, $s^*(M) \geq \hat{s}_c$. If the cut-off strategy of the voters $\hat{s} < s^*(M)$, then $L(\hat{s}) > 1$ and $p_{sim}(0,\beta) < 0$ from equation (3.A.2). If $\hat{s} > s^*(M)$, then $L(\hat{s}) < 1$ and $p_{sim}(0,\beta) > 0$ from equation (3.A.2). From (3.7) and (3.8), it follows that $\hat{s} = s_1(c)$, which contracts that $\hat{s} > s^*(M) \geq \hat{s}_c$. Finally, let the voters use the cut-off strategy $s^*(M)$. The organizer must choose $p_{sim}(M,\beta) = 1$ to satisfy (3.A.1) since $L(s^*(M)) = 1$. The strategy p_{sim} satisfies the optimality condition (3.7) since $F(s^*(M)|\beta) \geq c$. Hence, we construct a unique simple equilibrium.

Second, consider the case where $c > F(s^*(M)|\beta)$, that is, $\hat{s}_c > s^*(M)$. By (3.9), the voters must use the cut-off strategy $\hat{s} \geq \hat{s}_c$. If $\hat{s} > \hat{s}(c)$, then $L(\hat{s}) < 1$ and $p(0,\beta) > 0$ by (3.A.2), which contracts (3.7). Finally, let the voters use the cut-off strategy $\hat{s}_c$. From (3.A.2), the organizer must choose $p_{sim}(0,\beta) = 1 - L(\hat{s}_c)$ and $p_{sim}(M,\beta) = L(\hat{s}_c)$ to satisfy (3.A.1). The strategy p_{sim} satisfies the optimality condition (3.8) since $F(\hat{s}_c|\beta) = c$. Hence, we construct a unique simple equilibrium.

In Appendix 3.A.4, we analyze other responsive equilibria. Denote a responsive equilibrium as *interior equilibrium* if it is not a simple equilibrium. We show that, for each $c \in (0,1)$, there exists $M_{int}(c)$ such that for each $M > M_{int}(c)$, an interior equilibrium exists.

3.A.2 Proof of Proposition 3.4

For each $c \in (0,1)$, consider the sub-sequence of the voters

$$\mathbb{T}_c = \{t_c(1), ..., t_c(T)\} \subset \{1, ...M\}$$

such that

$$F(\hat{s}_{t_c(i)}|\alpha) > \frac{c}{2}, \forall i \in \{1, ..., T\}.$$

Note that $\mathbb{T}_c$ has at least one element. Otherwise, the organizer stops initially in state α.

Lemma 3.3. *For each $c \in (0,1)$, there exists T_c such that for each $M > 1$ and each non-trivial equilibrium, we have $t_c(1) < T_c$.*

Proof. We calculate one upper bound of the organizer's expected payoff in state α if he keeps sampling the voters. It is reached by letting the $t_c(1)^{th}$ voter use $\hat{s} = 1$ and letting the voters before her use $\hat{s}$ such that $F(\hat{s}|\alpha) = \frac{c}{2}$. Denote this upper-bound as $U(t_c(1))$,

$$U(t_c(1)) = \sum_{n=1}^{t_c(1)-1} \left(1 - \frac{c}{2}\right)^{n-1} (\frac{c}{2} - c) + \left(1 - \frac{c}{2}\right)^{t_c(1)-1} (1 - c).$$

Note that $U(t_c(1))$ is decreasing in $t_c(1)$ and

$$\lim_{t_c(1) \to \infty} U(t_c(1)) < 0.$$

Hence, we can choose

$$T_c = inf\{t_c(1) \in \mathbb{N} | U(t_c(1)) < 0\}.$$

□

Considering the organizer's sampling problem after the $t_c(i)^{th}$ voter rejects a for each $i \in \{1, ..., T\}$. Based on the proof of Lemma 3.3,

$$t_c(i+1) - t_c(i) \leq T_c, \ \forall i \in \{1, ..., T\}.$$

Hence,

$$T \geq \frac{M}{T_c}. \tag{3.A.3}$$

Now, consider a sequence of non-trivial equilibria with $M \to \infty$. By (3.A.3), we have $T \to \infty$. Therefore, the probability of implementing a in state α converges to 1.

In state β, for the i^{th} voter with $i \in \{1, ..., M\}$, the likelihood that she is sampled is

$$\left[1 - \sum_{j=0}^{i-1} p(j, \beta)\right] \cdot \prod_{j=1}^{i-1} \left[1 - F(\hat{s}_j|\beta)\right].$$

The first term is the probability that the organizer does not stop before reaching her. The second term is the probability that all the voters before her reject a. The i^{th} voter is pivotal if the voters sampled after her reject a. The likelihood is

$$\underbrace{\frac{p(i,\omega)}{1 - \sum_{k=0}^{i-1} p(k,\omega)}}_{\text{stop after the } i^{th} \text{ voter}} + \sum_{j=i+1}^{M} \left\{ \underbrace{\frac{p(j,\omega)}{1 - \sum_{k=0}^{i-1} p(k,\omega)}}_{\text{stop after the } j^{th} \text{ voter}} \cdot \prod_{k=i+1}^{j} \left[1 - F(\hat{s}_k|\beta)\right] \right\},$$

which is larger than

$$\prod_{j=i+1}^{M} \left[1 - F(\hat{s}_j|\beta)\right],$$

since

$$\prod_{k=i+1}^{M} \left[1 - F(\hat{s}_k|\beta)\right] \leq \prod_{k=i+1}^{j} \left[1 - F(\hat{s}_k|\beta)\right] \leq 1, \ \forall j \in \{i+1, ..., M\}.$$

Hence, the likelihood that the i^{th} voter is sampled and pivotal in state β is larger than

$$\left[1 - \sum_{j=0}^{i-1} p(j, \beta)\right] \cdot \prod_{j \neq i} \left[1 - F(\hat{s}_j|\beta)\right].$$

Denote $\Phi_i(s, \text{spl}, \text{piv})$ as the posterior likelihood ratio assigned by the i^{th} voter that the state is α, conditional on (i) receiving a signal s, (ii) being sampled, and (iii) being pivotal,

$$\Phi_i(s, \text{spl}, \text{piv}) \leq \frac{\pi}{1-\pi} \frac{f(s|\alpha)}{f(s|\beta)} \frac{\prod_{j \neq i} \left[1 - F(\hat{s}_j|\alpha)\right]}{\left[1 - \sum_{j=0}^{i-1} p(j, \beta)\right] \cdot \prod_{j \neq i} \left[1 - F(\hat{s}_j|\beta)\right]}.$$

We also have,

$$\begin{cases} \Phi_i(\hat{s}_i, \text{spl}, \text{piv}) = 1, \text{ if } \hat{s}_i \in (0,1), \\ \Phi_i(\hat{s}_i, \text{spl}, \text{piv}) \leq 1, \text{ if } \hat{s}_i = 0. \end{cases}$$

Therefore,

$$1 - \sum_{j=0}^{i-1} p(j, \beta) \leq \frac{\pi}{1-\pi} \frac{f(\hat{s}_i|\alpha)}{f(\hat{s}_i|\beta)} \frac{\prod_{j \neq i} \left[1 - F(\hat{s}_j|\alpha)\right]}{\prod_{j \neq i} \left[1 - F(\hat{s}_j|\beta)\right]}, \text{ if } \hat{s}_i \in (0,1).$$

Since $T \to \infty$ as $M \to \infty$,

$$\lim_{M\to\infty} \frac{\prod_{j\neq i}\left[1-F(\hat{s}_j|\alpha)\right]}{\prod_{j\neq i}\left[1-F(\hat{s}_j|\beta)\right]} = 0.$$

Hence, for each $\epsilon_1 > 0$ and $\epsilon_2 > 0$, we can find $M(\epsilon_1, \epsilon_2)$ such that for each $M > M(\epsilon_1, \epsilon_2)$, for each $i \in \{1, ..., M\}$, either

$$1 - \sum_{j=0}^{i-1} p(j,\beta) < \epsilon_1.$$

or[10]

$$\hat{s}_i < \epsilon_2.$$

Hence,

$$\lim_{M\to\infty} \hat{s}_i \left[1 - \sum_{j=0}^{i-1} p(j,\beta)\right] = 0. \tag{3.A.4}$$

Denote ψ_i as the probability that the organizer receives the approval in state β from the i^{th} voter,

$$\psi_i = \left[1 - \sum_{j=0}^{i-1} p(j,\beta)\right] \cdot \prod_{j=0}^{i-1}\left[1 - F(\hat{s}_j|\beta)\right] \cdot F(\hat{s}_i|\beta).$$

Note that the probability of implementing a in state β is $\sum_{i=1}^{M} \psi_i$.

From (3.A.4), for each $i \in \{1, ..., M\}$,

$$\lim_{M\to\infty} \psi_i = 0.$$

Since T_c is independent of M,

$$\lim_{M\to\infty} \sum_{i=1}^{T_c} \psi_i = 0.$$

We claim that

$$\lim_{M\to\infty} \sum_{i=T_c+1}^{M} \psi_i = 0.$$

Note that

$$\sum_{i=T_c+1}^{M} \psi_i < 1 - \sum_{j=0}^{T_c} p(j,\beta),$$

10. Note that we can replace this one by $\hat{s}_i = 0$ if $\lim_{s\to 0} \frac{f(s|\alpha)}{f(s|\beta)} < \infty$.

since if the organizer stops before reaching the $(T_c+1)^{th}$ voter, he cannot receives approval from her or the voters after her. Hence, it is sufficient to show that

$$\lim_{M\to\infty}\sum_{j=0}^{T_c} p(j,\beta) \;=\; 1. \tag{3.A.5}$$

If (3.A.5) does not hold, then we can construct a subsequence of non-trivial equilibria $\{\Gamma_{m_j}\}_{j=1,2,\ldots}$ such that

$$\lim_{m_j\to\infty}\sum_{j=0}^{n} p(j,\beta) \;<\; 1,\; \forall n \;\le\; T_c.$$

By (3.A.4),

$$\lim_{m_j\to\infty}\hat{s}_n \;=\; 0,\; \forall n \;\le\; T_c,$$

which contradicts Lemma 3.3.

Hence,

$$\lim_{M\to\infty}\sum_{i=1}^{M}\psi_i \;=\; 0.$$

The probability of implementing b in state β converges to 1 in each sequence of non-trivial equilibria with $M\to\infty$.

3.A.3 Set of Symmetric BNE

In this section, we characterize the set of symmetric Bayesian Nash equilibrium and show that we can focus on the cut-off strategy without loss of generality.

The strategy of the voters is

$$d\;:\;[0,1]\;\to\;[0,1].$$

A voter with signal $s\in[0,1]$ approves a with probability $d(s)$.

We denote $q(\omega)$ as the expected probability that one voter approves a in state $w\in\{\alpha,\beta\}$,

$$q(\omega)\;=\;\int_0^1 d(s)\mathrm{d}F(s|\omega).$$

The strategy of the organizer is still characterized by

$$p\;:\;\{0,\ldots,M\}\times\{\alpha,\beta\}\;\to\;[0,1],$$

such that

$$\sum_{n=0}^{M} p(n,\omega)\;=\;1,\;\forall\omega\;\in\;\{\alpha,\beta\}.$$

For each voter, when the strategy of the organizer is p and the strategies of other voters are d, the posterior likelihood ratio that the state is α, conditional on

(i) receiving a signal s, (ii) being sampled, and (iii) being pivotal, is denoted by $\Phi(s, \mathrm{spl}, \mathrm{piv}; p, d, M)$. We suppress arguments p, d, M and denote it by

$$\Phi(s, spl, piv) = \underbrace{\frac{\pi}{1-\pi}}_{\text{prior}} \underbrace{\frac{f(s|\alpha)}{f(s|\beta)}}_{\text{signal}} \underbrace{\frac{\sum_{m=1}^{M} p(m,\alpha)\ \frac{m}{M}\ (1-q(\alpha))^{m-1}}{\sum_{n=1}^{M} p(n,\beta)\ \frac{n}{M}\ (1-q(\beta))^{n-1}}}_{\text{sampled and pivotal under distribution } p}. \tag{3.A.6}$$

If the denominator is 0 but the numerator is not, then we set $\Phi(s, spl, piv) = \infty$. If both the denominator and the numerator are 0, then we set $\Phi(s, spl, piv) = 1$.

The strategy d is optimal if

$$\begin{cases} d(s) = 1 & \text{when } \Phi(s, spl, piv) > 1, \\ d(s) = 0 & \text{when } \Phi(s, spl, piv) < 1, \\ d(s) \in [0,1] & \text{when } \Phi(s, spl, piv) = 1. \end{cases} \tag{3.A.7}$$

The organizer's sampling problem is stationary. The strategy p is optimal if

$$\begin{cases} p(0,\omega) = 1 & \text{when } q(\omega) < c, \\ p(M,\omega) = 1 & \text{when } q(\omega) > c, \\ (p(0,\omega), \ldots., p(M,\omega)) \in \Delta\{[0,1]^{M+1}\} & \text{when } q(\omega) = c. \end{cases} \tag{3.A.8}$$

An equilibrium is a strategy profile $\{d, p\}$ satisfying (3.A.7) and (3.A.8).

We now show that we can focus on the cut-off strategy without loss of generality. We first consider the equilibrium where the organizer stops initially in both states. In such equilibrium, the voters use strategy d such that the corresponding approval probabilities $q(\omega) \le c$ for $\omega \in \{\alpha, \beta\}$. Note that we can construct such equilibrium by letting the voters use cut-off strategy $\hat{s}$ such that $F(\hat{s}|\alpha) \le c$. Furthermore, for the equilibrium in which all the voters always approve a with probability 1 when being sampled, the organizer keeps sampling the voters in both states since $c < 1$. In this equilibrium, the voters use cut-off strategy $\hat{s} = 1$.

Consider the equilibrium where (i) the organizer does active sampling in both states, and (ii) the voters do not always approve a. In such equilibrium, both the denominator and the numerator of (3.A.6) is positive. Hence, the posterior likelihood ratio $\Phi(s, spl, piv)$ is strictly increasing in s. The voters must use a cut-off strategy.

3.A.4 Set of Responsive Equilibria

In this section, we characterize the set of responsive equilibria and show that the simple equilibrium is the worst responsive equilibrium for the voters.

A responsive equilibrium is *complex* if

$$p(0,\beta) + p(M,\beta) < 1, \tag{3.A.9}$$

$$\sum_{n=o}^{M} p(n,\beta) \;=\; 1. \tag{3.A.10}$$

In a complex equilibrium $(p_\beta, \hat{s})$, since $p(M,\beta) < 1$, from (4), we have $\hat{s} = \hat{s}_c$. Plug in this to (16),

$$\Phi(\hat{s}_c, spl, piv) \;=\; \frac{\pi}{1-\pi}\,\frac{f(\hat{s}_c|\alpha)}{f(\hat{s}_c|\beta)}\,\frac{(1-F(\hat{s}_c|\alpha))^{M-1}}{\sum_{n=1}^{M} p(n,\beta)\ \frac{n}{M}\ (1-F(\hat{s}_c|\beta))^{n-1}} \;=\; 1. \tag{3.A.11}$$

Proposition 3.5. *For each $c \in (0,1)$, there exists $M_2(c)$ such that for each $M > M_2(c)$, a complex equilibrium exists.*

Proof. Fix the cost c. Since $\Phi(\hat{s}_c, spl, piv; p_\beta)$ is continuous with respect to p_β, if there exist vectors p'_β and p''_β satisfying (3.A.9) and (3.A.10) such that

$$\Phi(\hat{s}_c, spl, piv; p'_\beta) \;>\; 1, \tag{3.A.12}$$

$$\Phi(\hat{s}_c, spl, piv; p''_\beta) \;<\; 1, \tag{3.A.13}$$

then at least one p_β satisfying (3.A.9), (3.A.10) and (3.A.11). On the one hand, we can choose $p'_\beta = (1-\epsilon_M, \epsilon_M, 0, \ldots, 0)$ with ϵ_M near 0 to satisfy (3.A.12). On the other hand, Pick $p''_\beta = \{\frac{1}{2}, \frac{1}{2}, 0, \ldots, 0\}$. From (13),

$$\Phi(\hat{s}_c, spl, piv; p''_\beta) \;\leq\; \frac{\pi}{1-\pi}\,\frac{f(\hat{s}_c|\alpha)}{f(\hat{s}_c|\beta)}\,\frac{(1-F(\hat{s}_c|\alpha))^{M-1}}{\frac{1}{2}(1-F(\hat{s}_c|\beta))^{M-1}},$$

$$\lim_{M\to\infty} \Phi(\hat{s}_c, spl, piv; p''_\beta) \;\leq\; \lim_{M\to\infty} \frac{\pi}{1-\pi}\,\frac{f(\hat{s}_c|\alpha)}{f(\hat{s}_c|\beta)}\,\frac{(1-F(\hat{s}_c|\alpha))^{M-1}}{\frac{1}{2}(1-F(\hat{s}_c|\beta))^{M-1}} \;=\; 0.$$

Thus, we can pick $M_2(c)$ such that for each $M \geq M_2(c)$,

$$\Phi(\hat{s}_c, spl, piv; p''_\beta) \;<\; 1.$$

Therefore, for each $M \geq M_2(c)$, at least one complex equilibrium exists. □

Now, we prove that the simple equilibrium is the worst responsive equilibrium for the voters.

Proposition 3.6. *For each $c \in (0,1)$ and $M > 1$, the voters gain the lowest expected payoff from the simple equilibrium among all responsive equilibria.*

Proof. From proof of Proposition 3.1, we define define $s^*(M)$ by[11]

$$\frac{\pi}{1-\pi}\frac{f(s^*(M)|\alpha)}{f(s^*(M)|\beta)}\left[\frac{1-F(s^*(M)|\alpha)}{1-F(s^*(M)|\beta)}\right]^{M-1} = 1. \tag{3.A.14}$$

We already show that for the simple equilibrium, (i) $p(0,\beta)=1$ and $\hat{s}=s^*(M)$ if $c\le F(s^*(M)|\beta)$, and (ii) $p(0,\beta)<1$ and $\hat{s}=\hat{s}_c$ if $c>F(s^*(M)|\beta)$.

Case 1: $c\le F(s^*(M)|\beta)$

The expected payoff of one voter in the simple equilibrium is

$$\pi\{1-[1-F(s^*(M)|\alpha)]^M\}+(1-\pi)[1-F(s^*(M)|\beta)]^M.$$

We rewrite it by plugging in (3.A.14),

$$\pi+\pi[1-F(s^*(M)|\alpha)]^{M-1}\{[1-F(s^*(M)|\beta)]\frac{f(s^*(M)|\alpha)}{f(s^*(M)|\beta)}-1+F(s^*(M)|\alpha)\}. \tag{3.A.15}$$

The expected payoff of one voter in a complex equilibrium is

$$\pi\{1-[1-F(\hat{s}_c|\alpha)]^M\}+(1-\pi)\sum_{n=0}^{M}p(n,\beta)(1-F(\hat{s}_c|\beta))^n. \tag{3.A.16}$$

From (3.A.11),

$$(1-\pi)\sum_{n=1}^{M}p(n,\beta)(1-F(\hat{s}_c|\beta))^n > \pi\frac{f(\hat{s}_c|\alpha)}{f(\hat{s}_c|\beta)}[1-F(\hat{s}_c|\alpha)]^{M-1}(1-F(\hat{s}_c|\beta)).$$

Plug in this to (3.A.16), the expected payoff of one voter in a complex equilibrium is larger than

$$\pi+\pi[1-F(\hat{s}_c|\alpha)]^{M-1}\{[1-F(\hat{s}_c|\beta)]\frac{f(\hat{s}_c|\alpha)}{f(\hat{s}_c|\beta)}-1+F(\hat{s}_c|\alpha)\}. \tag{3.A.17}$$

Since $c\le F(s^*(M)|\beta)$, we have $\hat{s}_c\le s^*(M)$. By taking a derivative,

$$\frac{d(1-F(s|\alpha))}{ds} < 0,$$
$$\frac{d\{[1-F(s|\beta)]\frac{f(s|\alpha)}{f(s|\beta)}-1+F(s|\alpha)\}}{ds} < 0.$$

Therefore, when $c\le F(s^*(M)|\beta)$, the expected payoff from each complex equilibrium is higher than the one from the simple equilibrium.

Case 2: $c>F(s^*(M)|\beta)$

11. If it does not admit a solution, let $s^*(M)=0$.

In this case, the voters choose $\hat{s} = \hat{s}_c$ in each responsive equilibrium. The expected payoff of one voter is given by (3.A.16). Now, we solve the vector p_β minimizing (3.A.16), satisfying (3.A.10) and (3.A.11).

Define

$$\begin{aligned} \gamma &= 1 - F(\hat{s}_c|\beta), \\ K &= \pi \frac{f(\hat{s}_c|\alpha)}{f(\hat{s}_c|\beta)}[1 - F(\hat{s}_c|\alpha)]^{M-1}(1 - F(\hat{s}_c|\beta)), \end{aligned}$$

Note that only the second part of (3.A.16) depends on p_β. By writing (3.A.11) as a linear constrain in $\{p(0,\beta), \ldots, p(M,\beta)\}$, we are facing a linear programming problem:

$$\begin{aligned} \min_{\{p(0,\beta),\ldots,p(M,\beta)\}} \sum_{m=0}^{M} p(m,\beta)\gamma^m, & \\ \text{s.t} \quad \sum_{m=0}^{M} p(m,\beta) &= 1, \\ \sum_{m=1}^{M} p(m,\beta)\gamma^m m &= K, \\ p(m,\beta) &\geq 0, \qquad \forall m \in \{1,\ldots,M\}. \end{aligned}$$

Note that if the linear programming problem above admits one solution, then it admits one solution with at most two non-zero entries. Hence, the problem can be rewritten as

$$\min_{n_1,n_2\in\{0,\ldots,M\}} \quad p(n_1,\beta)\gamma^{n_1} + p(n_2,\beta)\gamma^{n_2}, \tag{3.A.18}$$

$$\text{s.t} \quad p(n_1,\beta) + p(n_2,\beta) = 1, \tag{3.A.19}$$

$$p(n_1,\beta)\gamma^{n_1}n_1 + p(n_2,\beta)\gamma^{n_2}n_2 = K, \tag{3.A.20}$$

$$p(n_1,\beta) \geq 0 \quad \text{and} \quad p(n_2,\beta) \geq 0. \tag{3.A.21}$$

If we can drop (3.A.21), then we can replace $p(n_1,\beta)$ and $p(n_2,\beta)$ in equation (3.A.18) by using (3.A.19) and (3.A.20), facing an unconstrained problem. Now, we try to find a way to drop (3.A.21).

Define m_c such that

$$m_c = \mathit{inf}\{n|n \in \{0,\ldots,M\}, \gamma^n n \geq K\}.$$

The existence of m_c is ensured by the existence of the responsive equilibrium. If m_c does not exist, then the linear programming does not admit one solution, which means there does not exist any responsive equilibrium.

Since $c > F(s^*(M)|\beta)$,

$$\frac{\pi}{1-\pi} \frac{f(\hat{s}_c|\alpha)}{f(\hat{s}_c|\beta)} \left[\frac{1-F(\hat{s}_c|\alpha)}{1-F(\hat{s}_c|\beta)}\right]^{M-1} < 1.$$

Therefore, we have $\gamma^M M > K$.

For any $\gamma \in (0,1)$, the function $\gamma^n n$ is single-peaked in n. We can divide $\{0,1,...,M\}$ to two sets $\{0,...,m_c - 1\}$ and $\{m_c,...,M\}$. If $n \in \{0,...,m_c - 1\}$, then $\gamma^n n < K$. If not, then $\gamma^n n \geq K$.

We can drop (3.A.21) by choosing $n_1 \in \{0,...,m_c - 1\}$ and $n_2 \in \{m_c,...,M\}$. Replace $p(n_1,\beta)$ and $p(n_2,\beta)$ in equation (3.A.18) by using (3.A.19) and (3.A.20). Finally, we are facing the following problem.

$$\min_{n_1 \in \{0,...,m_c-1\}, n_2 \in \{m_c,...,M\}} g(n_1,n_2) \tag{3.A.22}$$

where

$$g(n_1,n_2) \;=\; \frac{(n_2 - n_1)\gamma^{n_1+n_2} - K\gamma^{n_1} + K\gamma^{n_2}}{\gamma^{n_2} n_2 - \gamma^{n_1} n_1}.$$

Given $x \in [0, m_c - 1]$ and $y \in [m_c, M]$,

$$\frac{\partial g(x,y)}{\partial x} \;>\; 0,$$
$$\frac{\partial g(x,y)}{\partial y} \;<\; 0.$$

Thus, the solution of (3.A.22) and (3.A.18) is $n_1 = 0$ and $n_2 = M$. We can conclude that the simple equilibrium minimizes the expected payoff of each voter among all responsive equilibria when $c > F(s^*(M)|\beta)$. □

References

Ali, S Nageeb, and Navin Kartik. 2012. "Herding with collective preferences." *Economic Theory* 51 (3): 601–26. [87]

Banerjee, Abhijit V. 1992. "A simple model of herd behavior." *Quarterly Journal of Economics* 107 (3): 797–817. [86]

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch. 1992. "A theory of fads, fashion, custom, and cultural change as informational cascades." *Journal of Political Economy* 100 (5): 992–1026. [86]

Dekel, Eddie, and Michele Piccione. 2000. "Sequential voting procedures in symmetric binary elections." *Journal of Political Economy* 108 (1): 34–55. [76]

Duggan, John, and César Martinelli. 2001. "A Bayesian model of voting in juries." *Games and Economic Behavior* 37 (2): 259–94. [75, 77]

Feddersen, Timothy, and Wolfgang Pesendorfer. 1998. "Convicting the innocent: The inferiority of unanimous jury verdicts under strategic voting." *American Political Science Review* 92 (1): 23–35. [75]

Krishna, Vijay, and John Morgan. 2011. "Overcoming ideological bias in elections." *Journal of Political Economy* 119 (2): 183–211. [87]

Krishna, Vijay, and John Morgan. 2015. "Majority rule and utilitarian welfare." *American Economic Journal: Microeconomics* 7 (4): 339–75. [87]

Martinelli, César. 2006. "Would rational voters acquire costly information?" *Journal of Economic Theory* 129 (1): 225–51. [87]