Public Persuasion in Elections: Single-Crossing Property and the Optimality of Censorship*

Junze Sun[†] Arthur Schram[‡] Randolph Sloof[§]

This version: September 16, 2022. Click here for the latest version.

Abstract

We study public persuasion in elections with binary outcomes, like referendums. In our model, one or multiple senders attempt to influence the election outcome by manipulating public information about a payoff-relevant state. We allow for a wide class of sender preferences, ranging from pure self-interest to a broad set of social welfare functions, including utilitarian preferences. Our main result identifies a *single-crossing property* that ensures the optimality of *censorship policies*, which reveal intermediate states while censoring extreme states. This holds in large elections under both monopolistic and competitive persuasion. The single-crossing property holds for all self-interested senders and more generally under a mild condition for the distribution of voters' preferences. We analyze how a sender's optimal censorship policy changes with his preferences and voting rules, and characterize the asymptotic properties as the electorate size goes to infinity. Finally, our results shed new lights on whether media competition maximizes voter welfare.

JEL Codes: D72, D82, D83

Keywords: voting, single-crossing property, censorship, Bayesian persuasion, competition in persuasion, information design

*First version: February 8, 2022. We thank Zeinab Aboutalebi, Ricardo Alonso, Giacomo Calzolari, Piotr Dworczak, Federico Innocenti, Navin Kartik, Nenad Kos, David Levine, Elliot Lipnowski, Laurent Mathevet, Niko Matouschek, Salvatore Modica, Daniel Monte, Mallesh Pai, Marzena Rostek, Alex Smolin, Erik Snowberg, Jakub Steiner, Shoshana Vasserman, Asher Wolinsky, Takuro Yamashita, Andriy Zapechelnyuk, Jidong Zhou, and Zhen Zhou for helpful comments and discussions. We also thank audiences at the 5th International Conference on "The Political Economy of Democracy and Dictatorship" (PEDD) in Germany (virtual), the Young Economists' Meeting 2022 in Brno, the 12th Conference on Economic Design in Padova, the Max Weber program multidisciplinary research workshop on "Voters' Preferences and Parties' Performance in Politics" at the EUI, Econometric Society Summer Meetings (Asia, Australia), the 33rd Stony Brook International Conference on Game Theory, EEA-ESEM at Bocconi University, and seminar participants at the Nanyang Technological University and the European University Institute for their comments and feedback. Junze Sun gratefully acknowledges financial support from MIUR PRIN prj-0317_d1. All errors are our own.

[†]Corresponding author. Department of Economics, European University Institute and Department of Mathematics and Information, University of Palermo. Email: Junze.Sun@eui.eu.

[‡]Amsterdam School of Economics, University of Amsterdam. Email: A.J.H.C.Schram@uva.nl.

[§]Amsterdam School of Economics, University of Amsterdam. Email: R.Sloof@uva.nl.

1 Introduction

In modern democracies, many important choices are often made through collective decisions. For instance, presidents are selected via general elections and many important policies are determined in referendums. In general, many different individuals and organizations have diverse interests over the outcomes of such collective decisions: politicians, (possibly foreign) governments, mass media outlets, interest groups, representatives of industry or community leaders, etc. Anyone with a stake in the outcome may try to influence the election outcome through manipulating public information, e.g., via public announcements or debate.

This paper studies the strategic provision of public information in elections with binary outcomes, such as referendums. We model the environment of interest as a public Bayesian persuasion problem (Kamenica and Gentzkow, 2011), in which senders strategically choose public information policies to maximize their expected payoffs.[1] Our paper contributes to the literature (reviewed in Section 2) in two important dimensions. First, we allow for a broad set of utility functions for senders that embed both the pursuit of self-interest and maximizing the utilitarian welfare as special cases. Second, we characterize information provision in equilibrium under both monopolistic persuasion with a single sender and competitive persuasion with multiple senders. We implement both generalizations in a single, unified framework. Our central research question is: given the (possibly different) objectives of senders, what public information will be provided to voters in equilibrium?

Answering this question is important from both a positive and a normative perspective. From the positive view, it helps to understand the equilibrium behavior of actors interested in manipulating public information to influence election outcomes. From the normative view, our results shed light on the structure of the ideal public information policy for a social planner whose objective is to maximize a weighted average of voters' payoffs.

To illustrate our model (which is formally laid out in Section 3), consider a referendum where voters collectively decide between passing a reform and maintaining the status quo. An ex-ante unknown state k, which is drawn from a commonly known prior supported on a bounded interval, say $[-1,1]$, determines the quality of the reform relative to the status quo. Each voter is characterized by a private 'threshold of acceptance' v_i, such that her utility is $k-v_i$ and she prefers the reform if and only if $k \geq v_i$.[2] We refer to this threshold v_i as the voter's *type*. Voters' types are private information and are independently drawn from a commonly known distribution. Voters with higher type realizations receive lower ex-post payoffs if the reform is passed.

This setup fits many real-world scenarios in which the adoption of the reform can bring a public good of uncertain value, while at the same time induce idiosyncratic payoff shocks to voters. For

[1] Kamenica (2019) and Bergemann and Morris (2019) provide comprehensive overviews of this literature.
[2] Throughout this paper, we will refer to voters as feminine and senders as masculine.

example, consider a referendum on climate change in which the reform is a tax policy aimed at reducing emissions of greenhouse gasses, such as a car fuel levy or a tax on airline tickets.[3] The state k then represents the effectiveness of this tax policy in reducing emissions of greenhouse gases, which benefits the whole society. Voters' private types can, for instance, reflect the income shocks caused by the tax. The extent of such shocks depends on many idiosyncratic individual characteristics, such as a voter's occupation, employment status, wealth level, etc. Voters with higher types experience greater negative income shocks following the policy reform.[4]

There is a finite set of senders, who can provide voters with public information about state k. A sender can be anyone with an interest in and the ability to manipulate voters' public information. At the same time, when interested in normative implications we also view a sender as an abstract social planner who wants to maximize voter welfare. Like voters, each sender's utility is linear in state k and he prefers the reform if and only if k is above some threshold, say ϕ (which can differ across senders). For a *self-interested* sender who does not take voters' welfare into account, his threshold ϕ is independent of voters' types. Alternatively, a sender can also be *prosocial*; in this case he cares about voters' welfare and hence his threshold ϕ will depend on voters' private types. For example, a utilitarian planner's threshold ϕ equals the voters' average type; he prefers the reform if and only if voters are on average better off under the reform than under the status quo. More generally, we allow each sender's utility function to be any convex combination of his self-interest and some weighted average of all voters' payoffs. This allows for a broad spectrum of preferences.

Before knowing the realizations of either the state or voters' types, each sender simultaneously chooses an *information policy*, which maps any state realization k to a (distribution of) public signal. After observing their private types and the public signals jointly sent by the senders, voters simultaneously decide to vote for either the reform or the status quo. The reform will be adopted if and only if the fraction of votes it receives exceeds a cutoff set by the voting rule. For example, under the simple majority rule this cutoff is 50%. Since information transmission is public and voters' payoffs are linear in state, they must share the same posterior expectation about the state realization (which is sufficient to determine their voting behavior and expected payoffs).

[3] Such referendums have been held in practice. For instance, in June 2021 the Swiss People's Party launched a referendum on the Federal Act on the Reduction of Greenhouse Gas Emissions (CO2 Act). The goal of this act is to reduce emissions of carbon dioxide and other greenhouse gases in Switzerland by 50% (compared to 1990 levels) by 2030, using mainly tax policies. See https://en.wikipedia.org/w/index.php?title=2021_Swiss_referendums&oldid=1100537027.

[4] Our model also fits many other contexts beyond referendums. For example, many papers in the political economics literature adopt similar models to ours in studying the electoral competition between two politicians (Groseclose, 2001; Ashworth and De Mesquita, 2009; Chakraborty and Ghosh, 2016; Chakraborty, Ghosh and Roy, 2020; Alonso and Câmara, 2016*b*; Sun, Schram and Sloof, 2021). Here, the state can be interpreted as candidates' valences or competences, which are commonly appreciated by all voters. Each voter's private type is interpreted as her idiosyncratic ideology. Another example is committee voting by a board of directors in a company. The board members collectively decide whether to invest in a project, whose profit is the ex-ante unknown state. Private types measure each board member's reservation value.

One class of information policies that will prove to be particularly important are the so-called *censorship policies*, which have a simple interval-revelation structure as illustrated in Figure 1. With this policy, a sender will precisely reveal the realized state k if it lies in interval $[a,b]$, but only report "$k < a$" if the realization is below a and report "$k > b$" if the realization is above b. It is in this sense that state realizations outside of the revelation interval $[a,b]$ are 'censored'. Under such a policy, voters' posterior expected state equals the realized state k whenever it lies within the revelation interval $[a,b]$, while it equals $\mathbb{E}[k|k > b]$ for $k > b$ and $\mathbb{E}[k|k < a]$ for $k < a$.

Figure 1: Censorship Policy

The main result of our paper is the following. We identify a sufficient condition, which is widely held and has an appealing economic interpretation, that ensures that it is optimal for a sender to focus on censorship policies of the kind described in Figure 1. This is true under both monopolistic persuasion with a single sender and competitive persuasion with multiple sender.

Our sufficient condition can be interpreted as a *single-crossing property* over the sender's and the pivotal voter's *indifference curves*, which are derived as follows. Under any cutoff voting rule the election outcome is essentially determined by the choice of the *pivotal voter*, whose realized type is denoted by x.[5] The pivotal voter prefers the reform to the status quo if and only if $k \geq x$. Her indifference is therefore the 45-degree line on a two-dimensional plane with horizontal axis the pivotal voter's realized type x, and the vertical axis the realized state k. Now we draw a sender's indifference curve on the same plane. This task is straightforward for a self-interested sender; his indifference curve is simply a horizontal line in this plane, because his threshold of acceptance for the reform is independent of all voters' types (which of course include the pivotal voter's type x). Deriving the indifference curve for a prosocial sender is more subtle. The key tension here is that a prosocial sender's preference over the election outcome depends on voters' private types, which are unobservable to him. The sender must therefore *infer* his preference by exploiting the *statistical correlations* between the pivotal voter's type and the types of other voters. For example, suppose that the sender is a utilitarian social planner and the election outcome is determined by simple majority rule. Then, given any realized type profile of voters, the planner's threshold of acceptance for the reform is given by the average type (denoted by $\bar{v}$) while the pivotal voter's threshold of acceptance is the median type (denoted by v^m). Therefore, conditional on the pivotal voter's type being x, the sender rationally infers his expected threshold of acceptance to be $\mathbb{E}[\bar{v}|v^m = x]$, whose value

[5] For example, under simple majority rule the pivotal voter is the median voter, whose type x equals the sample median of voters' realized types.

depends on x, the distribution of voter's types, and the electorate size. This gives his indifference curve for all possible type realizations x of the pivotal voter. Such an inference procedure similarly applies to general social preferences and voting rules. The wedge between the indifference curves of the pivotal voter and a sender determines their conflict of interests, which is critical in shaping the sender's optimal information policy.

This brings us to our single-crossing property. Informally speaking (in Section 4 we present the formal definition), the single-crossing property holds for a sender if, in sufficiently large elections, his indifference curve crosses the pivotal voter's indifference curve at most once, and if so only from above. We show that the single-crossing property holds under very broad conditions. It is always satisfied if the sender is self-interested, and satisfied for all sender preferences and voting rules under a mild assumption on the distribution of voter preferences.

In Section 5 we analyze monopolistic persuasion by a single sender for whom the single-crossing property holds. In this case we show that some censorship policy as in Figure 1 must be uniquely optimal for this sender in sufficiently large elections (Theorem 1).[6] The optimal choices of the boundaries a and b are driven by the tradeoff between the capability of manipulating voters' beliefs in more states on the one hand (providing incentives to censor more states), and the effectiveness of belief manipulation on the other hand (reducing censoring incentives). In Section 6 we further characterize and discuss properties of the sender's optimal censorship policy and payoff as the electorate size goes to infinity (Theorem 2). We also derive comparative statics regarding how the optimal censorship policy varies with the sender's preference and the voting rule. Among other things, we find a novel *sender-preference effect*: ceteris paribus, increasing the required vote share to pass the reform makes any prosocial sender more inclined towards passing the reform. This effect stems from the statistical inference problem explained above, and it is absent when the sender is either (i) completely self-interested, or (ii) perfectly observes voters' private types (like in Alonso and Câmara (2016*a*)). We show that this effect generates novel implications for how a prosocial sender's optimal information policy should respond to changes in voting rules.

In Section 7 we study competitive persuasion where multiple senders simultaneously choose their public information policies as in Gentzkow and Kamenica (2017*b*). In this case we show that if the single-crossing property holds for a sender and the electorate size is sufficiently large, then it is *without loss of optimality* for this sender to focus on a subset of censorship policies in the following sense: for any feasible pure strategy profile chosen by other senders (which need not be censorship policies), this sender can always find a censorship policy from this subset as his best response (Theorem 3). Now suppose that the single-crossing property holds for all senders. Then, under a weak regularity condition, in the minimally informative equilibrium the information

[6] In Theorem 1 we also give conditions under which censorship policies are uniquely optimal independent of the electorate size. Hence, when these conditions hold, our results apply to small-size elections such as in committee voting.

jointly provided by all senders can be reproduced by a censorship policy, whose revelation interval is simply the convex hull of the revelation intervals that would be optimal for each of the senders under monopolistic persuasion (Theorem 4). We also characterize a sufficient condition under which competition in persuasion must induce full information revelation in all equilibria.

We finally apply our results to study the welfare implications of media competition. In our model, competition between two partisan and opposite-minded media outlets induces full information revelation in any equilibrium. Nevertheless, and perhaps surprisingly, such full disclosure is in general suboptimal from the welfare perspective. We compare voters' utilitarian welfare under the information policy that maximizes utilitarian welfare and under full information disclosure, as the electorate size goes to infinity. We show that the former is always larger than the latter, and the gap can be substantial if the ex-ante conflict of interests between the average voter and the pivotal voter is large. These results imply that it is important to account for the distribution of voters' preferences and voting rules – which jointly determine the ex-ante conflict of interests between the average and pivotal voters – when evaluating the welfare effects of media competition.

2 Related literature

This paper speaks to several strands of literature. First of all, our paper contributes to the literature that studies information transmission in elections using the Bayesian persuasion or information design approach.[7] Aside from a few exceptions discussed below, most papers in this literature study persuasion by a monopoly sender whose goal is to sway the election outcome in favor of his preferred alternative (Wang, 2013; Alonso and Câmara, 2016*a*,*b*; Bardhi and Guo, 2018; Chan et al., 2019; Ginzburg, 2019; Kerman, Herings and Karos, 2020; Heese and Lauermann, 2021; Gitmez and Molavi, 2022; Gradwohl, Heller and Hillman, 2022).[8] To our best knowledge, we are the first to provide a unified framework to study public Bayesian persuasion in elections in an environment that simultaneously allows for (i) a broad set of sender preferences, (ii) an arbitrary number of senders, and (iii) any super-majority voting rules.

The two studies closest to ours are Alonso and Câmara (2016*b*) and Kolotilin, Mylovanov and Zapechelnyuk (2022). The models in both papers can be interpreted as a monopoly sender

[7] Of course, strategic information transmission in elections has been extensively studied under various other communication protocols, such as cheap talk (Schnakenberg, 2015, 2017; Kartik and Van Weelden, 2019; Sun, Schram and Sloof, 2021) and verifiable disclosure (Liu, 2019). One important feature that separates our paper from these is that we can also address the normative question regarding the optimal information policy for a social planner.

[8] Some of these papers (e.g., Heese and Lauermann (2021)) allow the sender's preferred alternative to be state-dependent. They do not, however, allow for the sender's utility to depend on voters' payoffs. Moreover, all these papers except Alonso and Câmara (2016*a*,*b*) and Ginzburg (2019) study targeted persuasion in which the sender can privately communicate to voters (Bergemann and Morris, 2016; Taneva, 2019; Mathevet, Perego and Taneva, 2020). Our paper instead focuses on public persuasion whereby a sender must send the same message to all voters.

persuading a privately informed representative voter. In Alonso and Câmara (2016*b*), the sender is an incumbent party leader who aims at maximizing the re-election probability. They show that, under some regularity conditions, the optimal information policy is upper censorship if the distribution of the representative voter's private type has a log-concave density. Kolotilin, Mylovanov and Zapechelnyuk (2022) characterize sufficient and necessary conditions for the optimality of upper censorship for general linear persuasion problems. They show that the same log-concavity density assumption ensures this optimality for a wider class of sender preferences, ranging from maximizing the winning probability to maximizing the payoff of the representative voter.

Our paper enriches and generalizes the results of both papers to an environment that allows for multiple senders and voters. Looking at a setup with multiple voters instead of a single representative voter enables us to model a much broader class of social preferences for senders, and to study the influence of voting rules on the optimal information policy. We show that in large elections the optimality of censorship can be ensured under much weaker assumptions regarding the underlying distribution of voter types than those made in previous studies. Moreover, we establish that the same conditions that ensure the optimality of censorship for a sender under monopolistic persuasion continue to do so under competitive persuasion with multiple senders.

Alonso and Câmara (2016*a*) study public persuasion in elections by a monopoly sender and how the voting rule affects the optimal information policy in a model similar to ours. A crucial difference between our paper and theirs is that we allow voters to have private types, while in their model the sender knows voters' preferences. This difference is important in two ways. First, the structures of the optimal information policies are very different depending on whether voters' preferences are known to the sender. Second, we show that when a sender cares about social welfare and is imperfectly informed about voters' preferences, varying the voting rule can affect his optimal information policy through a novel sender-preference effect. This effect is absent if the sender has perfect information about voters. Van der Straeten and Yamashita (2020) and Ferguson (2020) study monopolistic persuasion problems for a utilitarian planner in models different from ours. Both papers show that full information disclosure is suboptimal from the utilitarian perspective. Our paper extends this insight to general social welfare functions. Finally, Innocenti (2021) and Mylovanov and Zapechelnyuk (2021) study competition in Bayesian persuasion by two opposite-minded senders with pure persuasion motives. The former does so in a model where each voter can only hear from one sender. The latter, like ours, consider public persuasion a la Gentzkow and Kamenica (2017*b*). Our paper allows for a much richer set of sender preferences compared to theirs.

Second, methodologically, our paper relates to a recent strand of literature that develops the duality approach to solve linear persuasion problems in which senders' utility functions depend only on the posterior expected state (Kolotilin, 2018; Dworczak and Martini, 2019; Dworczak and Kolotilin, 2019; Dizdar and Kováč, 2020; Kolotilin, Mylovanov and Zapechelnyuk, 2022;

Sun, 2022*a*,*b*).[9] In particular, Dworczak and Martini (2019) show that the problem of finding an equilibrium outcome under competitive persuasion can be converted to solving the monopolistic persuasion problems of each sender with modified utility functions. This allows us to treat monopolistic and competitive persuasion in a unified framework. Kolotilin, Mylovanov and Zapechelnyuk (2022) exploit the duality method to show that upper (resp. lower) censorship policies are uniquely optimal if the sender's utility function is strictly S-shaped (resp. inverse S-shaped) in posterior expectation. Sun (2022*b*) extends this observation to competitve persuasion in an environment a la Gentzkow and Kamenica (2017*b*); he shows that if a sender's utility function is strictly S-shaped (resp. inverse S-shaped), then given any pure strategy profile of others, there exists an upper (resp. lower) censorship policy as the sender's best response. Sun (2022*a*) uses the duality method to derive an easy-to-check sufficient condition for full information revelation under competition in persuasion in linear persuasion games. We build on these findings to establish our main results.

Finally, our results also relate to papers studying competition in Bayesian persuasion with multiple senders (Gentzkow and Kamenica, 2017*b*,*a*; Cui and Ravindran, 2020; Au and Kawai, 2020, 2021; Li and Norman, 2021; Mylovanov and Zapechelnyuk, 2021; Sun, 2022*a*). An important theme of this literature is to identify conditions under which full information disclosure is the unique equilibrium outcome. We contribute to this research agenda by providing such a sufficient condition in the context of publicly persuading voters. In contrast to many earlier works but consistent with Sun (2022*a*), we show that strong conflicts of interests between competing senders are not necessary to induce full information disclosure as the unique equilibrium outcome.

3 Framework

We consider an election in which $n+1$ voters collectively decide between two options, which for ease of reference we label *Reform* and *Status quo*. The outcome is determined by a cutoff rule with threshold $q \in (0,1)$; the reform is adopted if and only if it obtains strictly more than nq votes. For instance, $q = 0.5$ corresponds to simple majority rule. For ease of exposure, we assume that nq is an integer (unless explicitly mentioned otherwise).

An ex-ante unknown but payoff relevant state k is drawn from a common prior F that admits a positive and continuous density f on $[-1,1]$. Without loss of generality, we normalize all players' payoffs to zero under the status quo. If the reform is adopted, each voter i's payoff equals $k - v_i$, where v_i is her private type. In this way, voter i's payoff attributed to the reform (relative to the status

[9] Several papers study linear persuasion problems using other methods. For instance, Gentzkow and Kamenica (2016) and Kolotilin et al. (2017) characterize the sets of implementable outcomes under public and private signals, respectively, using the Blackwell theorem. More recently, Arieli et al. (2022), Ivanov (2020) and Kleiner, Moldovanu and Strack (2021) develop methods based on theories of extreme points and majorization to characterize structures of solutions to linear persuasion problems.

quo) consist of a common value k (the 'quality' of the reform), and her private type v_i. Because voter i prefers the reform if and only if $k \geq v_i$, her private type represents her threshold of acceptance of the reform. We assume that each v_i is independently drawn from a commonly known distribution G, which admits a positive and twice continuously differentiable density g on $[\underline{v}, \overline{v}]$ with $\underline{v} < -1$ and $\overline{v} > 1$. For any profile of type realizations $v = (v_1, \cdots, v_{n+1})$, we let $v^{(1)} \leq v^{(2)} \leq \cdots \leq v^{(n+1)}$ be its ascending permutation. Since $k - v_i$ decreases in v_i, voters with lower type realizations receive higher ex-post payoffs if the reform is adopted.

Consider first a monopoly sender; in Section 7 we extend our model to allow for multiple senders competing in persuading voters. The sender's payoff under the reform is given by

$$u(k, v) = \rho \sum_{j=1}^{n+1} w_j \cdot \left(k - v^{(j)}\right) + (1 - \rho)(k - \chi) \tag{1}$$

where $\rho \in [0, 1]$, $\chi \in \mathbb{R}$ and $(w_1, \cdots, w_{n+1})$ is a non-negative vector of weights that sum up to 1. Parameter ρ captures the extent to which the sender cares about 'voter welfare' relative to his 'self-interest'. If $\rho = 0$ then the sender prefers reform to be adopted if and only if $k \geq \chi$. In this case, the sender is *self-interested* in the sense that his preference over alternatives is independent of voters' interests.[10] Conversely, if $\rho = 1$ then $u(k, v)$ is a weighted average of voters' realized payoffs when the reform is adopted. For each $j = 1, \cdots, n+1$, w_j is the *rank-dependent welfare weight* the sender assigns to the voter whose payoff under reform is ranked the j-th highest under the realized type profile v. The vector $(w_1, \cdots, w_{n+1})$ is generated by a *weighting function* $w(\cdot)$ that is non-decreasing, absolutely continuous on $[0, 1]$ and satisfies $w(0) = 0$ and $w(1) = 1$. Hence, $w(\cdot)$ is the cumulative distribution function (cdf) of a random variable on $[0, 1]$.[11] For any integer $n \geq 0$ and $j \in \{1, \cdots, n+1\}$, element w_j is uniquely generated by

$$w_j = w\left(\frac{j}{n+1}\right) - w\left(\frac{j-1}{n+1}\right) \tag{2}$$

This setup captures a wide class of social welfare functions in a unified way. For instance, the utilitarian welfare function can be obtained by letting $\rho = 1$ and $w(x) = x$ for all $x \in [0, 1]$. With this $w(\cdot)$, it follows from (2) that $w_j = \frac{1}{n+1}$ for each j so that the welfare weights are equal across voters. If $w(\cdot)$ is not the cdf of a uniform distribution on $[0, 1]$, then it represents the preference of some non-utilitarian social planner who may discriminate voters according to the ranking of their

[10] This captures transparent persuasion motives (which are most extensively explored in the literature) as limiting cases. For instance, the preference of a sender whose aim is to maximize the winning probability of reform (resp. status quo) independent of state realizations can be captured by letting $\chi \to -\infty$ (resp. $\chi \to \infty$).

[11] $w(\cdot)$ is reminiscent of the probability weighting function in the rank-dependent utility theory (Quiggin, 1982). As we will see in Section 6.2, there is a natural connection between the first order stochastic dominance ordering of $w(\cdot)$ and the sender's social preference.

ex-post payoffs. We will discuss some examples in Section 5.

The sender can affect voters' information about k by designing an *information policy*. Following the convention of the Bayesian persuasion literature, we define an information policy π by a pair (S,σ), where S is a sufficiently rich signal space and $\sigma : [-1,1] \mapsto \Delta(S)$ maps each state realization k to a probability distribution on S. Let Π denote the set of all feasible information policies.

The timing of the game is as follows. First, prior to observing state k, the sender chooses an information policy $\pi \in \Pi$. Second, state k is realized and a public signal is drawn according to π. Observing the realized public signal, voters simultaneously decide to vote for either the reform or the status quo. The reform is adopted if and only if its vote tally strictly exceeds nq. All players' payoffs then realize. Throughout, we focus on equilibria in weakly undominated strategies.[12]

3.1 Voting behavior and election outcome

Because voters have a common prior F and information transmission is public, they must share a common posterior about the state realization after hearing from the sender. Since voters' payoffs under reform are linear in state k, their expected payoffs depend only on their posterior expectation θ and are given by $\theta - v_i$ for all i. It is then a weakly dominant strategy for voter i to vote for reform if and only if $\theta \geq v_i$. Therefore, under the cutoff voting rule with threshold q, the election outcome is determined by the choice of the *pivotal* voter, whose type realization is $v^{(nq+1)}$. Note that $v^{(nq+1)}$ is a random variable and let $\hat{G}_n(\cdot;q)$ denote its cumulative distribution function. Since reform is adopted only if $v^{(nq+1)} \leq \theta$, $\hat{G}_n(\theta;q)$ gives the winning probability of reform. Appendix A offers a formal expression and useful properties of $\hat{G}_n(\theta,q)$.

Lemma 1. *$\hat{G}_n(\cdot;q)$ is strictly increasing. $v^{(nq+1)}$ converges in probability to $v_q^* := G^{-1}(q)$.*[13]

Lemma 1 says that the winning probability of reform strictly increases in θ. Moreover, as $n \to \infty$ the reform is adopted with probability one (zero) if $\theta > (<) v_q^*$.

4 Indifference curves and the single-crossing property

In this section we introduce the single-crossing property and discuss its implications for a sender's temptation to manipulate voters' beliefs. We also characterize sufficient conditions for our single-crossing property. All derivations and proofs are in Appendix B.

[12] It is well known that the voting game at the second stage has a plethora of uninteresting equilibria in weakly dominated strategies. For example, whenever $n > 0$ it is an equilibrium for all voters to vote for reform regardless of their private types or the public information they obtain, because no single vote can unilaterally change the outcome. In this case, any information policy π can be sustained in equilibrium as well because they have no influence. The restriction to weakly undominated strategies rules out such uninteresting equilibria.

[13] Here, $G^{-1}(\cdot)$ is the inverse function of G, which is well defined because G is strictly increasing on its support $[\underline{v},\overline{v}]$. v_q^* is then the quantile function of a random variable with distribution G for each $q \in [0,1]$.

4.1 Indifference curves and the inference from pivotal voter's choice

Given any realization of voter type profile v, it follows from equation (1) that the sender weakly prefers the reform if and only if

$$k \geq \varphi_n(v) := \rho \sum_{j=1}^{n+1} w_j \cdot v^{(j)} + (1-\rho)\chi \; .$$

$\varphi_n(v)$ is the sender's threshold of acceptance for the reform. Note that this depends on voters' realized type profile v whenever $\rho > 0$. Importantly, however, at the time of choosing his information policy, any sender with $\rho > 0$ cannot precisely observe $\varphi_n(v)$ because realized types are voters' private information. Nevertheless, the election outcome, which is essentially the choice of the pivotal voter, is informative about the realization of $\varphi_n(v)$.

To make this point clear, it is instructive to draw the indifference curves of the pivotal voter and the sender in the same plane, as in Figure 2. In each panel, the horizontal axis x represents the pivotal voter's type realization $v^{(nq+1)}$ and the vertical axis denotes the realized state k. The pivotal voter's indifference curve is simply the 45-degree line; she is indifferent between alternatives if and only if $k = x$. Let

$$\phi_n(x) := \mathbb{E}\left[\varphi_n(v)\middle|v^{(nq+1)} = x\right] = \rho \sum_{j=1}^{n+1} w_j \cdot \mathbb{E}\left[v^{(j)}\middle|v^{(nq+1)} = x\right] + (1-\rho)\chi \qquad (3)$$

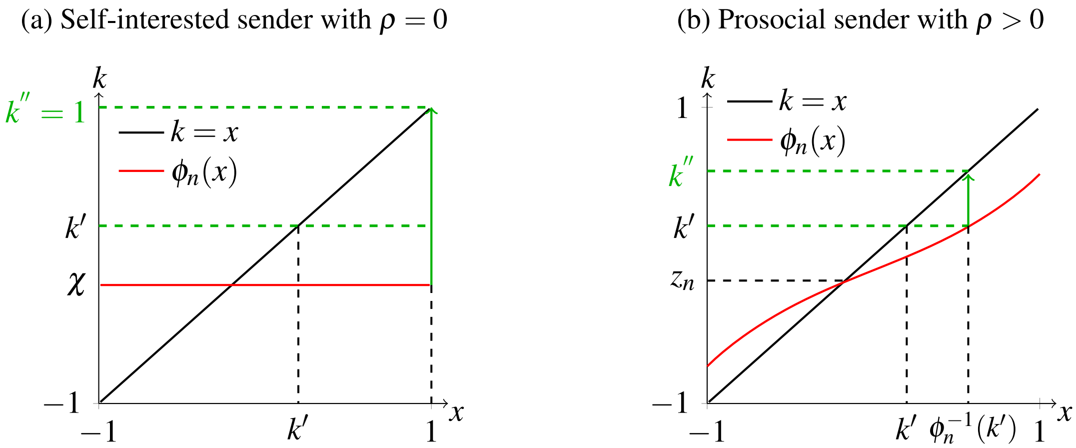
denote the expectation of $\varphi_n(v)$ conditional on event $v^{(nq+1)} = x$. Then, if the sender only knows that $v^{(nq+1)} = x$, he would be indifferent between alternatives if and only if $k = \phi_n(x)$. For this reason, we refer to $\phi_n(x)$ as the sender's *indifference curve*.

Panel (a) of Figure 2 depicts the indifference curve of a self-interested sender with $\rho = 0$. In this case it is obvious from (3) that $\phi_n(x) = \chi$ for all x. The preference of a self-interested sender is thus independent of the pivotal voter's type realization.

Panel (b) of Figure 2 depicts the indifference curve of a prosocial sender with $\rho > 0$. In this case, we show in Appendix B (cf. Proposition B.2) that $\phi_n(x)$ is strictly increasing in x for all $n \geq 0$ and weighting functions $w(\cdot)$. This is because the pivotal voter's type realization $v^{(nq+1)}$ is positively associated with all other order statistics $v^{(j)}$ for $j = 1, \cdots, n+1$. Therefore, no matter how the sender assigns his welfare weights, the pivotal voter's type realization is either *directly relevant* or *indirectly informative* about the sender's threshold of acceptance for the reform. It is in this way that the inference from the pivotal voter's choice is important for any prosocial sender.

Two remarks are in place. First, the inference problem here is conceptually different from the inference about the state conditional on the event of being pivotal, which is central to the literature on information aggregation in voting (e.g., Feddersen and Pesendorfer (1996, 1997)). In our model

Figure 2: Indifference Curves and the Single-Crossing Property



Note: In both panels the horizontal axis x denotes the pivotal voter's type realization $v^{(nq+1)}$, the black line denotes the pivotal voter's indifference curve, and the red line denotes the sender's indifference curve.

voters have no private information about state k, so the information aggregation issue is absent. Second, for our inference problem to be relevant, it is necessary that voters' types are their private information. Therefore, our inference problem disappears in models where the sender has complete information about voters' preferences, such as Alonso and Câmara (2016*a*).

4.2 The single-crossing property and its economic implications

To formally define our *single-crossing property* we need the following lemma, which characterizes the limit of $\phi_n(\cdot)$ as $n \to \infty$.

Lemma 2. *For $x \in [\underline{v}, \overline{v}]$, define*

$$\phi(x) := \rho\left[\int_0^q G^{-1}\left(\frac{y}{q}G(x)\right)dw(y) + \int_q^1 G^{-1}\left(\frac{y-q}{1-q} + \frac{1-y}{1-q}G(x)\right)dw(y)\right] + (1-\rho)\chi \quad (4)$$

As $n \to \infty$, $\phi_n(x)$ and its partial derivative $\phi_n'(x)$ converge uniformly to $\phi(x)$ and $\phi'(x)$, respectively, on $[\underline{v}, \overline{v}]$. Moreover, $\varphi_n(v)$ converges almost surely to

$$\phi^* := \phi(v_q^*) = \rho\int_0^1 G^{-1}(y)dw(y) + (1-\rho)\chi \quad (5)$$

For any continuously differentiable function $h(\cdot)$, we say that $h(\cdot)$ is *single-crossing* on $[-1,1]$ if (i) $h(x)$ crosses zero at most once and if so from below on $[-1,1]$, and (ii) $h'(x) > 0$ whenever $h(x) = 0$ and $x \in [-1,1]$.

Definition 1. *We say that the* single-crossing property *holds for a sender if $x - \phi(x)$ is single-crossing on $[-1,1]$.*

By Lemma 2, $\phi_n(x)$ and $\phi_n'(x)$ converge uniformly to $\phi(x)$ and $\phi'(x)$, respectively, on $[\underline{v},\overline{v}]$ (which contains $[-1,1]$ by assumption). Therefore, when the single-crossing property holds for the sender, there exists a threshold $\tilde{n}$ such that for all $n \geq \tilde{n}$ function $x-\phi_n(x)$ is single-crossing on $[-1,1]$; that is, $\phi_n(x)$ crosses the pivotal voter's indifference curve $k=x$ at most once and if so only from above (cf. Figure 2). For such $\phi_n(\cdot)$ we can define a unique *switching state* z_n as follows:

$$z_n := \begin{cases} -1 & \text{if } x > \phi_n(x) \text{ for all } x \in [-1,1] \\ x & \text{if } x = \phi_n(x) \text{ for some } x \in [-1,1] \\ 1 & \text{if } x < \phi_n(x) \text{ for all } x \in [-1,1] \end{cases}. \tag{6}$$

The definition of z_n implies $k > \phi_n(k)$ for $k > z_n$ and $k < \phi_n(k)$ for $k < z_n$. Therefore, in any state $k > z_n$ the sender is more biased towards the reform than the pivotal voter in the following sense: whenever the pivotal voter is indifferent (i.e., in event $k=x$) the sender must strictly prefer the reform to be adopted because $k > \phi_n(k)$. Similarly, in any state $k < z_n$ the sender is more biased towards the status quo than the pivotal voter in that he must strictly prefer the status quo to be maintained whenever the pivotal voter is indifferent.

An important economic implication of the single-crossing property is that the sender is tempted to manipulate voters' beliefs upwards (downwards) for state realizations above (below) the switching state z_n. Figure 2 illustrates this. Consider any state realization k' in $(z_n,1)$. Under the single-crossing property $k' > \phi_n(k')$ must hold. Let $k'' = \phi_n^{-1}(k')$ if $k' \leq \phi_n(1)$ (right panel) or set $k''=1$ otherwise (left panel). As is evident in Figure 2, $k'' > k'$ must hold so that the sender and pivotal voter prefer different alternatives whenever $x \in (k',k'')$, with the sender strictly preferring the reform. The sender is then tempted to lie and let the pivotal voter believe that the realized state is k'', which is higher than the true state k'. It is in this sense that the sender is tempted to manipulate voters' beliefs about state upwards. Following the same logic, if the state realization k' is below z_n then the sender is tempted to manipulate voters' beliefs about the state downwards.

4.3 Sufficient conditions for the single-crossing property

This subsection provides two easy-to-check sufficient conditions for the single-crossing property. By Definition 1, the single-crossing property is ensured if $\phi'(\cdot) < 1$, that is, the sender's indifference curve is 'uniformly flatter' than the pivotal voter's indifference curve. Lemma 3 provides two sufficient conditions for this to hold.

Lemma 3. *Suppose either (i) ρ is sufficiently close to* 0*, or (ii) both G and $1-G$ are strictly log-concave.*[14] *Then $\phi'(\cdot) < 1$, and $\phi_n'(\cdot) \leq 1$ for all $n \geq 0$.*

[14] Strict log-concavity of $1-G$ is equivalent to a strictly increasing hazard rate $g(x)/(1-G(x))$. Strict log-concavity

Conditions (i) simply says that the sender is sufficiently self-interested. Condition (ii) requires very mild conditions on the distribution of voter preferences. These are satisfied if the density function g is strictly log-concave, which already includes a wide class of distributions (see Bagnoli and Bergstrom (2005) for examples) that are frequently assumed in applied theories.[15] Once this mild assumption for G is satisfied, the single-crossing property holds for a sender regardless of his preferences and the voting rules.

5 Main result: The single-crossing property and the optimality of censorship policy

This section presents our main result, which relates the single-crossing property to the optimality of censorship policies for a monopoly sender in sufficiently large elections (Section 5.1). A formal presentation of the persuasion problem and the proof of the main result are in Section 5.2.

5.1 Optimal information policy under monopolistic persuasion

Consider a monopoly sender and let $\phi_n(x)$ be his indifference curve. We assume that the single-crossing property holds, so there exists $\tilde{n} \geq 0$ such that for all $n \geq \tilde{n}$ function $x - \phi_n(x)$ is single-crossing on $[-1,1]$ and the unique switching state z_n is identified by (6).

As explained in the Introduction, a censorship policy is characterized by a revelation interval $[a,b]$ with $-1 \leq a \leq b \leq 1$ such that (i) all intermediate state realizations $k \in [a,b]$ are precisely revealed, and (ii) extreme state realizations $k > b$ and $k < a$ are censored under different pooling messages as in Figure 1. Under a censorship policy, voters' (common) posterior expectation equals k for all state realizations $k \in [a,b]$ due to full revelation, and equals $\mathbb{E}_F[k|k > b]$ (resp. $\mathbb{E}_F[k|k < a]$) for all state realizations $k > b$ (resp. $k < a$). Observe that both *full disclosure* (with $a = -1$ and $b = 1$) and *no disclosure* (with $a = b \in \{-1,1\}$) are special cases of censorship policies.

Our main result, Theorem 1, relates the single-crossing property to the optimality of censorship policies in large elections under monopolistic persuasion.

Theorem 1. *Consider a monopoly sender for whom the single-crossing property holds. Then there exists an $N \geq 0$ such that for all $n \geq N$ any optimal information policy is outcome equivalent to a*

of G is equivalent to a strictly decreasing reversed hazard rate $g(x)/G(x)$. In fact, this condition is tight; suppose that either G or $1-G$ is strictly log-convex on some sub-interval within $[-1,1]$, then it is possible to construct a sender preference and voting rule under which the single-crossing property fails to hold.

[15] Our condition for G is in fact significantly less demanding than the strict log-concavity of g in the following sense. g must be single-peaked and hence G is uni-model if g is strictly log-concave. This, however, need not be the case under condition (ii) in Lemma 3. Therefore, our condition for G allows for some polarized distributions which put more probability weights on extreme states than on intermediate ones.

censorship policy with revelation interval $[a_n, b_n]$ *that satisfies* $-1 \le a_n \le z_n \le b_n \le 1$.[16] *Moreover, the following holds:*

1. *If* $-1 < z_n < 1$ *then* $a_n < z_n < b_n$ *so that the revelation interval contains the switching state* z_n *in its interior.*

2. *If* $z_n = -1$*, then* $a_n = -1$ *so that only sufficiently high states can be censored.*

3. *If* $z_n = 1$*, then* $b_n = 1$ *so that only sufficiently low states can be censored.*

In addition, if $g(\cdot)$ *is strictly log-concave and* ρ *is sufficiently close to* 0*, then* $N = 0$ *so that these three properties hold for all* $n \ge 0$*.*

Theorem 1 establishes a one-to-one mapping between the three possible locations of the switching state z_n and the structure of the optimal censorship policy. If $z_n \in (-1, 1)$ so that $\phi_n(x) - x$ crosses zero at some interior state, then the optimal policy has the feature of *two-sided censorship* in the sense that both very high and very low states can be censored. If instead $z_n = -1$, then $\phi_n(x) < x$ for all $x \in (-1, 1)$ and the sender is uniformly more biased towards reform than the pivotal voter. In this case the optimal policy takes the form of *upper censorship* in the sense that only sufficiently high states can be censored. Finally, if $z_n = 1$, then $\phi_n(x) > x$ for all $x \in (-1, 1)$ and the sender is uniformly more biased towards the status quo than the pivotal voter. In this case the optimal policy takes the form of *lower censorship* in that only sufficiently low states can be censored. Observe that Theorem 1 is robust in the sense that it applies for all continuous prior F, and for all G as long as the single-crossing property holds. By Lemma 3, this implies that Theorem 1 holds for generic G if the sender is sufficiently self-interested,[17] and it holds for all sender preferences characterized by (1) if both G and $1 - G$ are strictly log-concave.

The intuition underlying Theorem 1 is as follows. Observe that when $\phi_n(x) - x$ crosses zero from above at an interior switching state $z_n \in (-1, 1)$, the revelation interval $[a_n, b_n]$ of the optimal censorship policy must contain z_n in its interior so that voters can always perfectly distinguish between state realizations above and below z_n. Indeed, the single-crossing property implies that the sender has no incentive to hide state realization $k = z_n$. This is because at the switching state $k = z_n$ the interests of the sender and the pivotal voter are aligned; whenever the pivotal voter strictly prefers either alternative, the sender weakly prefers it. Moreover, it is always optimal for the sender

[16] Two information policies are *outcome equivalent* if their induced mappings from state realization k to voters' posterior expected state are equal almost everywhere.

[17] The result would be sharply different in a setup with only one representative voter (i.e., $n = 0$). This case is studied by Kolotilin, Mylovanov and Zapechelnyuk (2022). They show that G must be uni-model (i.e., g is single-peaked) to ensure the optimality of a censorship policy for a self-interested sender with $\rho = 0$. In our setup this is not required because, as Proposition A.1 in Appendix A shows, the density function of the pivotal voter's type distribution $\hat{g}_n(\cdot; q)$ is single-peaked for sufficiently large n for all g that are positive and twice-continuously differentiable.

to fully separate any pair of state realizations on different sides of z_n. To see why, consider any k_1 and k_2 such that $k_n < z_n < k_2$ and suppose they are not fully separated. As explained above, the sender is tempted to manipulate voters' beliefs about state realizations upwards in state k_2 while downwards in k_1. By fully separating these two states, the induced posterior expectation about state realization will indeed be lower in k_1 and higher in k_2 than in any other case where they are not fully separated. The sender thus strictly benefits from such separation.

What, then, drives the optimal choices of thresholds a_n and b_n? Consider b_n first. To perfectly separate state realizations above and below z_n, $b_n \geq z_n$ must hold. Now, recall that the sender is tempted to manipulate voters' beliefs upwards for all states $k > z_n$. Suppose the sender increases threshold b_n to some $b_n + \Delta$ with $\Delta > 0$ small. Then the sender loses the opportunity to manipulate voter's beliefs upwards in any state $k \in [b_n, b_n + \Delta]$ because these states are now fully revealed. Nevertheless, this expansion of b_n increases the induced posterior expectation from $\mathbb{E}_F[k|k > b_n]$ to $\mathbb{E}_F[k|k > b_n + \Delta]$ – so that the pivotal voter is more likely to be convinced to pass the reform – in all states $k \in [b_n + \Delta, 1]$. The optimal choice of b_n therefore balances the marginal costs of losing the capability to manipulate voters' beliefs in some states with the marginal gains of a increased effectiveness of persuasion. The tradeoff governing the optimal choice of a_n is similar.

Now we apply Theorem 1 to illustrate the structures of the optimal censorship policies for four examples of sender preferences. These are demonstrated by the four panels of Figure 3.

Example 1. Self-interested sender. *Panel (a) depicts the indifference curve and structure of the optimal censorship policy for a self-interested sender with $\rho = 0$. By (3), $\phi_n(x) = \chi$ for all $x \in [\underline{v}, \overline{v}]$. His switching state z_n thus depends solely on χ. If $\chi \in (-1, 1)$, then $z_n = \chi$ and by Theorem 1 some two-sided censorship policy with $a_n < \chi < b_n$ is optimal for this sender in large elections. If instead $\chi \leq -1$ (resp. $\chi \geq 1$), then he is uniformly more biased towards the reform (resp. status quo) than the pivotal voter in all states. For these cases, Theorem 1 implies the sender's optimal information policy must be either upper (if $\chi \leq -1$) or lower (if $\chi \geq 1$) censorship in large elections.*

Example 2. Utilitarian social planner. *Panel (b) considers the case of a Utilitarian planner who aims at maximizing voters' ex-post average payoffs. His indifference curve $\phi_n(x)$ is given by*[18]

$$\phi_n(x) = \frac{n}{n+1}\left(q\mathbb{E}_G[v_i|v_i \leq x] + (1-q)\,\mathbb{E}_G[v_i|v_i \geq x]\right) + \frac{1}{n+1}x \tag{7}$$

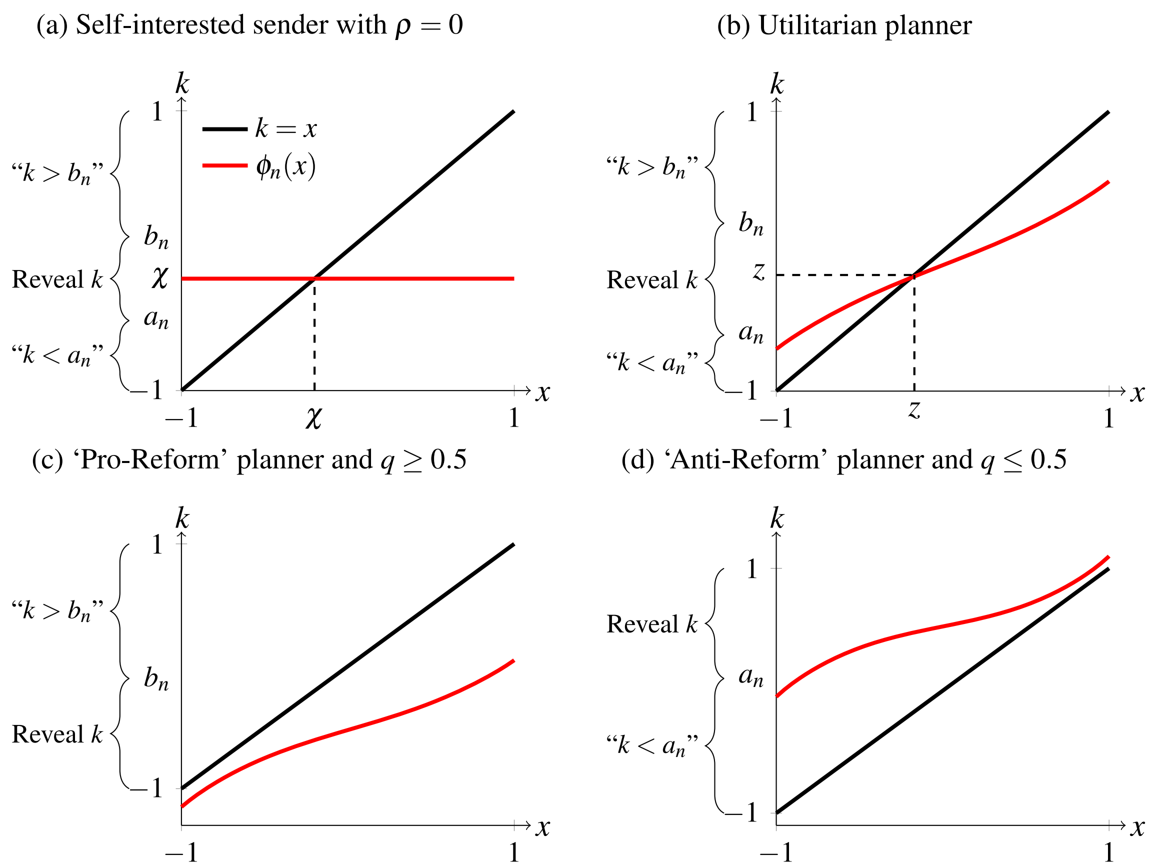
for $x \in [\underline{v}, \overline{v}]$. Therefore, $\phi_n(x) = x$ if and only if

$$q\mathbb{E}_G[v_i|v_i \leq x] + (1-q)\,\mathbb{E}_G[v_i|v_i \geq x] = x \tag{8}$$

[18] To see why (7) is true, notice that for event $v^{(nq+1)} = x$ to hold, there must be one voter with type $v_i = x$, nq other voters with $v_i \leq x$, and the remaining $n(1-q)$ voters with $v_i \geq x$. Since each voter's type is independently drawn from G, the conditional expectation of any voter with $v_i \leq x$ (resp. $v_i \geq x$) equals $\mathbb{E}_G[v_i|v_i \leq x]$ (resp. $\mathbb{E}_G[v_i|v_i \geq x]$). Taking the average over the whole electorate size $n+1$ yields (7).

When both G and $1-G$ are strictly log-concave, the single crossing property holds by Lemma 3 and hence (8) admits a unique solution z on $(\underline{v}, \overline{v})$. A Utilitarian planner's switching point z_n thus depends only on z. If $z \in (-1,1)$ as depicted in panel (b), then $z_n = z$ and by Theorem 1 some two-sided censorship policy with $a_n < z < b_n$ is Utilitarian optimal in large elections. Interestingly, if $z \leq -1$ (resp. $z \geq 1$) then even a Utilitarian planner can be uniformly more biased towards reform (resp. status quo) than the pivotal voter. For these cases the Utilitarian optimal information policy is either upper (if $z \leq -1$) or lower (if $z \geq 1$) censorship in large elections.

Figure 3: The Structures of Optimal Censorship Policies in Four Examples



Note: In each of these panels the horizontal axis x denotes the pivotal voter's type realization $v^{(nq+1)}$, the black line denotes the pivotal voter's indifference curve, and the red line denotes the sender's indifference curve.

Example 3. 'Pro-Reform' social planner. *In panel (c) we consider a non-Utilitarian social planner who aims at maximizing the average payoff of the subset of voters whose ex-post payoffs under reform are above the 50% percentile.*[19] *Suppose $q \geq 0.5$ so that a strict majority is required in order to pass the reform. In this case, $\phi_n(x) < x$ must hold for all $x \in (-1,1)$; that is, the sender must be uniformly more biased towards the reform than the pivotal voter in all states. This is*

[19] The weighting function for such a 'pro-Reform' planner is given by $w(x) = \min\{2x, 1\}$ for $x \in [0,1]$.

because he assigns positive weights only to voters who always like the reform better than the pivotal voter does. The sender thus must prefer the reform if the pivotal voter is indifferent. Therefore, by Theorem 1, in large elections some upper censorship policy must be optimal.

Example 4. 'Anti-Reform' social planner. *In panel (d) we consider a non-Utilitarian social planner who aims at maximizing the average payoff of the subset of voters whose ex-post payoffs under reform are below the 50% percentile.*[20] *Following the same logic as in Example 3, we can show that for all $q \leq 0.5$ (i.e., a strict majority is required to maintain the status quo) $\phi_n(x) > x$ must hold for all $x \in (-1,1)$; that is, such a sender must be uniformly more biased towards the status quo than the pivotal voter in all states. Theorem 1 then implies that some lower censorship policy must be optimal in large elections.*

We conclude this subsection with two remarks. First, if the sender can perfectly observe voters' types $(v_1, \cdots, v_{n+1})$, our model then becomes a special case of Alonso and Câmara (2016*a*) and the optimal information policy is a binary cutoff strategy that only reveals whether the realized k is above or below some threshold. Therefore, the fact that our optimal information policy typically has a more nuanced structure (i.e., with a non-degenerate revelation interval) is due to the assumption that voters have private types. Second, as Examples 2 to 4 illustrate, full information disclosure can be suboptimal even when the sender's goal is to maximize voters' welfare. To understand why, observe that conditional on the pivotal voter's type realization being x, in all states k between x and $\phi_n(x)$ the preferences of the sender and pivotal voters disagree. Our single-crossing property implies that $x = \phi_n(x)$ can hold for at most one $x \in [-1,1]$ for sufficiently large n. The interim conflict of interests between the sender and the pivotal voter is thus ubiquitous. Consequently, full disclosure is in general suboptimal.

5.2 The persuasion problem and proof of Theorem 1

We start by formally presenting the persuasion problem faced by the monopoly sender. Recall θ denotes the common posterior expectation about the state realization, shared by all voters and the sender. Given $\phi_n(\cdot)$, the sender's (indirect) expected utility under θ is

$$W_n(\theta) = \int_{\underline{v}}^{\theta} (\theta - \phi_n(x))\, \hat{g}_n(x;q)dx \tag{9}$$

where $\hat{g}_n(\cdot;q) = \hat{G}'_n(\cdot;q)$ is the density function of the pivotal voter's type realization. Because the sender's expected payoff depends on voters' posterior expectation θ only, it is convenient to present any information policy π by the distribution H_π of posterior means it induces. We say that a

[20] The weighting function for such an 'anti-Reform' planner is given by $w(x) = \max\{2x-1, 0\}$ for $x \in [0,1]$.

distribution of posterior means H is *feasible* if it can be induced by some information policy $\pi \in \Pi$. It is well known that given prior F, a distribution of posterior means H is feasible if and only if F is a *mean-preserving spread* of H (Gentzkow and Kamenica, 2016; Kolotilin et al., 2017; Dworczak and Martini, 2019).[21] In the sequel we write $F \succeq_{MPS} H$ if F is a mean-preserving spread of H.

The persuasion problem. For a monopoly sender, an information policy π is optimal if H_π solves

$$\max_{H\in\Delta([-1,1])} \int_{-1}^{1} W_n(\theta)dH(\theta), \quad \text{s.t. } F \succeq_{MPS} H \tag{MP}$$

As we show in Appendix C.2, $W_n(\cdot)$ is twice-continuously differentiable and thus upper semi-continuous. Therefore, (MP) admits at least one solution (Dworczak and Martini, 2019).

Finally, observe that the distribution of the posterior expectation induced by a censorship policy with revelation interval $[a,b]$ is given by

$$H_{\mathscr{P}(a,b)}(\theta) := \begin{cases} F(a)\cdot \mathbb{1}\{\theta \geq \mathbb{E}_F[k|k<a]\}, & \text{if } \theta \in [-1,a) \\ F(\theta), & \text{if } \theta \in [a,b] \\ F(b) + [1-F(b)]\cdot \mathbb{1}\{\theta \geq \mathbb{E}_F[k|k>b]\}, & \text{if } \theta \in [b,1] \end{cases} \tag{10}$$

where $\mathbb{1}\{\cdot\}$ is the indicator function. We say that an information policy $\pi \in \Pi$ is a *censorship policy* if H_π coincides with (10) for some $-1 \leq a \leq b \leq 1$ almost everywhere. In the sequel we slightly abuse notation and let $\mathscr{P}(a,b)$ denote both any specific censorship policy or the set of all censorship policies with revelation interval $[a,b]$, whenever this does not lead to confusion. For the special case $a = b$ we simply write $\mathscr{P}(a,b)$ as $\mathscr{P}(a)$ and refer to it as a *cutoff policy* because it only reveals whether the realize state is above, equal, or below cutoff a.

Our proof for Theorem 1 relies on the following two lemmas, whose proofs are in Appendix C. Both lemmas establish important curvature properties of $W_n(\cdot)$ that help to pin down the structures of solutions to the sender's problem (MP). In Lemma 4, we identify a novel '*increasing slope property*' and show that this condition ensures that any solution H to (MP) cannot be less informative than a cutoff policy that only reveals whether the realized state is above, equal or below a certain threshold.

Lemma 4. *Suppose that $x - \phi_n(x)$ crosses zero only once and from below at an interior point $z_n \in (-1,1)$. Then $W_n(\cdot)$ satisfies the 'increasing-slope property' at point z_n, that is,*

$$\frac{W_n(x) - W_n(z_n)}{x - z_n} \leq \frac{W_n(y) - W_n(z_n)}{y - z_n}, \forall y > x$$

[21] F is a mean-preserving spread of H if $\int_{-1}^{x} H(\theta)d\theta \leq \int_{-1}^{x} F(\theta)d\theta$ for all $x \in [-1,1]$ and the equality holds for $x = \pm 1$. An alternative, equivalent definition is that $\mathbb{E}_F[\omega(\cdot)] \geq \mathbb{E}_H[\omega(\cdot)]$ for any convex function $\omega(\cdot)$.

and strict inequality holds if $x < z_n < y$ (cf. panel (a) of Figure 4).[22] *Moreover, any solution H to problem (MP) must satisfy $H \succeq_{MPS} H_{\mathscr{P}(z_n)}$. In other words, any optimal information policy must reveal whether the state realization is above, equal to or below z_n.*

In line with standard terminology (e.g., Kolotilin, Mylovanov and Zapechelnyuk (2022)), we say that $W_n(\cdot)$ is *strictly S-shaped* on some interval if it is strictly convex below some inflection point and strictly concave above it. Likewise, $W_n(\cdot)$ is *strictly inverse-S-shaped* if it is strictly concave below some inflection point and strictly concave above it. Notice that both definitions include strictly convex and concave functions as special cases.

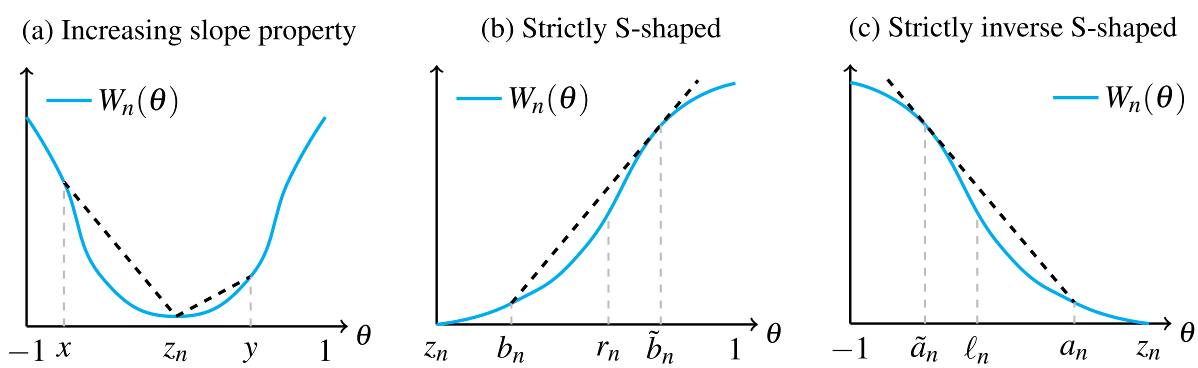
Lemma 5. *Suppose that the single-crossing property holds. Then there exists an $N \geq 0$ such that for all $n \geq N$ there are ℓ_n and r_n with $-1 \leq \ell_n \leq z_n \leq r_n \leq 1$ such that:*[23]

1. $W_n(\cdot)$ is strictly S-shaped on $[z_n, 1]$ with inflection point r_n (cf. panel (b) of Figure 4); and

2. $W_n(\cdot)$ is strictly inverse-S-shaped on $[-1, z_n]$ with inflection point ℓ_n (cf. panel (c) of Figure 4).

In addition, if $g(\cdot)$ is strictly log-concave and ρ is sufficiently close to 0, then $N = 0$ so the above curvature properties hold for all $n \geq 0$.

Figure 4: Graphical illustrations of Lemmas 4, 5 and the proof of Theorem 1



With these ingredients we are ready to prove Theorem 1, directly using the quantities N, ℓ_n and r_n identified in Lemma 5.

Proof of Theorem 1. Depending on the value of z_n, we distinguish between three cases.

Case 1: $\phi_n(x) - x$ crosses zero from above at a unique interior point $z_n \in (-1, 1)$. By Lemma 4, $W_n(\cdot)$ satisfies the increasing-slope property at point z_n and any solution H to problem (MP)

[22] Geometrically, a function $U(\cdot)$ satisfies the increasing-slope property at point z only if for all $x \neq z$ the line segment connecting $(x, U(x))$ and $(z, U(z))$ lies above $U(\cdot)$, as demonstrated in panel (a) of Figure 4.

[23] In fact, single-crossing property is almost necessary; suppose instead that $\phi(x) - x$ crosses zero from below at some point, then this lemma no longer holds and for sufficiently large n there exists some interval $[x, y] \subset (-1, 1)$ and $\varepsilon > 0$ such that $W_n(\cdot)$ is strictly concave on $[x, y]$ but is strictly convex on $[x - \varepsilon, x]$ and $[y, y + \varepsilon]$, respectively. In this case, it follows from duality arguments in Dworczak and Martini (2019) and Kolotilin, Mylovanov and Zapechelnyuk (2022) that there exists some continuous and full-support prior F under which the optimal information policy is not censorship.

must satisfy $H \succeq_{MPS} H_{\mathscr{P}(z_n)}$. As a consequence, the monopolistic persuasion problem can be decomposed into two auxiliary problems on intervals $[-1, z_n]$ and $[z_n, 1]$, respectively:

$$\max_{H \in \Delta([z_n,1])} \int_{z_n}^{1} W_n(\theta) dH(\theta), \quad \text{s.t. } F_{\text{I}} \succeq_{MPS} H \tag{MP-I}$$

$$\max_{H \in \Delta([-1,z_n])} \int_{-1}^{z_n} W_n(\theta) dH(\theta), \quad \text{s.t. } F_{\text{II}} \succeq_{MPS} H \tag{MP-II}$$

In these problems, F_{I} is the truncated cdf of F on $[z_n, 1]$, and it equals F if $z_n = -1$. F_{II} is the truncated cdf of F on $[-1, z_n]$, and it equals F if $z_n = 1$.

Recall from Lemma 5 that, for all $n \geq N$, $W_n(\cdot)$ is strictly S-shaped on $[z_n, 1]$ with inflection point r_n and strictly inverse S-shaped on $[-1, z_n]$ with inflection point ℓ_n. Then, by Kolotilin, Mylovanov and Zapechelnyuk (2022), the solution to problem (MP-I) is uniquely given by a censorship policy $\mathscr{P}(z_n, b_n)$. The threshold b_n satisfies the following complementary slackness condition

$$(\tilde{b}_n - b_n) W_n'(\tilde{b}_n) \leq W_n(\tilde{b}_n) - W_n(b_n) \tag{FOC: b_n}$$

where $\tilde{b}_n = \mathbb{E}_F[k | k \geq b_n]$ and (FOC: b_n) is binding whenever $b_n \in (z_n, 1)$ (cf. panel (b) of Figure 4).[24] Moreover, b_n and $\tilde{b}_n$ satisfy $b_n < r_n < \tilde{b}_n$ for $r_n \in (z_n, 1)$, and $b_n = 1$ if $r_n = 1$. Similarly, the solution to problem (MP-II) is uniquely given by a censorship policy $\mathscr{P}(a_n, z_n)$. The threshold a_n satisfies the following complementary slackness condition

$$(a_n - \tilde{a}_n) W_n'(\tilde{a}_n) \leq W_n(a_n) - W_n(\tilde{a}_n) \tag{FOC: a_n}$$

where $\tilde{a}_n = \mathbb{E}_F[k | k \leq a_n]$ and (FOC: a_n) is binding whenever $a_n \in (-1, z_n)$ (cf. panel (c) of Figure 4). Moreover, a_n and $\tilde{a}_n$ satisfy $a_n > \ell_n > \tilde{a}_n$ for $\ell_n \in (-1, z_n)$, and $a_n = -1$ if $\ell_n = 1$. Taken together, these imply that the optimal solution is uniquely given by a censorship policy $\mathscr{P}(a_n, b_n)$.

Next we show that $a_n < z_n < b_n$ must hold whenever $z_n \in (-1, 1)$. By (9) we have

$$\begin{aligned} W_n''(z_n) &= \hat{g}_n(z_n; q)\left(2 - \phi_n{}'(z_n)\right) + (z_n - \phi_n(z_n))\, \hat{g}_n'(z_n; q) \\ &= \hat{g}_n(z_n; q)\left(2 - \phi_n{}'(z_n)\right) > \hat{g}_n(z_n; q) > 0 \end{aligned}$$

The second step holds because $z_n - \phi_n(z_n) = 0$ by definition of z_n, and the third step holds because the single-crossing property requires $\phi_n'(z_n) < 1$ whenever $z_n - \phi_n(z_n) = 0$. Therefore, $W_n(\theta)$ is strictly convex in a neighborhood around z_n and thus $r_n > z_n$. This implies that

$$W_n'(z_n) < \frac{W_n(\theta) - W_n(z_n)}{\theta - z_n} \tag{11}$$

[24] If $(\tilde{b}_n - b_n) W_n'(\tilde{b}_n) > W_n(\tilde{b}_n) - W_n(b_n)$ for all $b_n \in [z_n, 1]$ then $b_n = 1$. The similar result holds for a_n.

holds for all $\theta \in [z_n, r_n]$. Since $W_n(\cdot)$ satisfies the increasing-slope property at point z_n, the right-hand side of (11) is increasing in θ. Therefore, (11) must hold for all $\theta > z_n$. It follows directly that condition (FOC: b_n) cannot be binding at $b_n = z_n$ for any $z_n \in (-1,1)$. This implies that $b_n > z_n$ must hold. $a_n < z_n$ can be established analogously. This proves statement (1) of Theorem 1.

Case 2: $z_n = -1$ so that $\phi_n(x) < x$ for all $x \in (-1,1)$. By Lemma 5, $W_n(\cdot)$ is S-shaped on $[-1,1]$ with inflection point $r_n \in [-1,1]$ for all $n \geq N$. The optimal information policy is therefore uniquely given by an upper-censorship policy $\mathscr{P}(-1, b_n)$, with b_n determined by condition (FOC: b_n). This proves statement (2) of Theorem 1.

Case 3: $z_n = 1$ so that $\phi_n(x) > x$ for all $x \in (-1,1)$. By Lemma 5, $W_n(\cdot)$ is strictly inverse S-shaped on $[-1,1]$ with inflection point $\ell_n \in [-1,1]$ for all $n \geq N$. The optimal information policy is therefore uniquely given by a lower-censorship policy $\mathscr{P}(a_n, 1)$, with a_n determined by condition (FOC: a_n). This proves statement (3) of Theorem 1. □

6 Properties of the optimal censorship policy and the sender's payoff

Our main result Theorem 1 shows that if the single-crossing property holds for a sender, then there exists a threshold $N \geq 0$ such that for all $n \geq N$ a censorship policy with revelation interval $[a_n, b_n]$ is uniquely optimal under monopolistic persuasion. In this section we further explore the properties of these optimal thresholds. In Section 6.1, we characterize and discuss properties of a_n and b_n and the sender's payoff as the electorate size $n \to \infty$. In Section 6.2 we study how the optimal thresholds a_n and b_n vary with sender preferences and voting rules.

6.1 The asymptotically optimal censorship policy and the sender's payoff

In this subsection we study the asymptotically optimal censorship policy and the sender's payoff as $n \to \infty$. In doing so we address important questions. For instance, under what conditions will the sender opt for some partially informative policy that features neither full disclosure nor no disclosure? What are conditions under which the sender's optimal policy censors states both upwards and downwards? We also derive the sender's asymptotic payoff under his optimal information policy and compare it to the 'first-best' and full information benchmarks. Such questions are of interest because they provide further intuition for the drivers of censorship policies and their consequences for the sender. Omitted proofs for this subsection are in Appendix D.

Our main result for this subsection is Theorem 2, which shows that under the single-crossing property the asymptotically optimal censorship policy can be characterized with only three variables:

$v^*_q = G^{-1}(q)$, $\phi^* = \phi(v^*) = \rho\int_0^1 G^{-1}(y)dw(y) + (1-\rho)\chi$, and

$$z^* := \lim_{n\to\infty} z_n = \begin{cases} -1 & \text{if } x > \phi(x) \text{ for all } x \in [-1,1] \\ x & \text{if } x = \phi(x) \text{ for some } x \in [-1,1] \\ 1 & \text{if } x < \phi(x) \text{ for all } x \in [-1,1] \end{cases} . \tag{12}$$

Before formally stating Theorem 2, let us recall the economic implications of the three variables. By Lemmas 1 and 2 we have $v^{(nq+1)} \xrightarrow{p} v^*_q$ and $\varphi_n(v) \xrightarrow{a.s.} \phi^*$, respectively. Therefore, as $n \to \infty$, the pivotal voter prefers reform (status quo) almost surely if $k > (<) v^*_q$, while the sender prefers reform (status quo) almost surely if $k > (<) \phi^*$. The difference between v^*_q and ϕ^* thus reflects the *ex-ante* conflict of interests between the sender and the pivotal voter, i.e. *before the actual type of the pivotal voter is drawn*. Finally, z^* is the limit of switching states z_n and, following the discussion in Section 4.2, for all states $k > z^*$ (resp. $k < z^*$) the sender is tempted to manipulate voters' posterior expectation about k upwards (resp. downwards) as $n \to \infty$. This reflects the effect of the inference problem from the pivotal voter's choice explained in Section 4.1. Unlike the ex-ante conflict of interests, this inference procedure determines the wedge between preferences of the pivotal voter and the sender *conditional on each possible realization of the pivotal voter's type*. Importantly, z^* can differ from ϕ^* only when the sender is uncertain about his preference. In our model this can happen only if both (i) voters have private information regarding their preferences, and (ii) $\rho > 0$ so that the sender cares about voters' payoffs. Lemma 6 gives important qualitative properties on the relationship between the three variables.

Lemma 6. *Suppose $\phi^* \in (-1,1)$. Then v^*_q, ϕ^* and z^* satisfy the following properties:*

1. *If $\phi^* = v^*_q$, then $z^* = \phi^* = v^*_q$.*
2. *If $\phi^* \neq v^*_q$, then $z^* = \phi^*$ if $\rho = 0$, and either $z^* < \phi^* < v^*_q$ or $z^* > \phi^* > v^*_q$ if $\rho > 0$.*

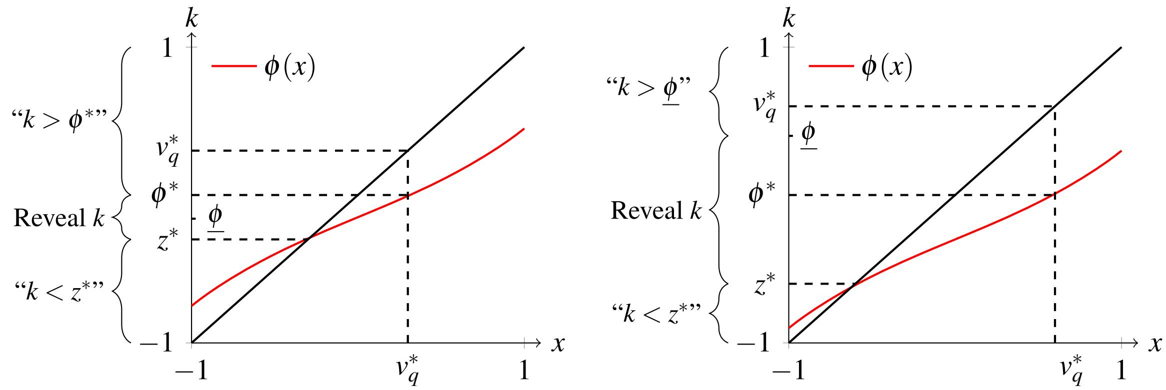
Figure 5 illustrates Lemma 6 for the case $\rho > 0$ and $\phi^* < v^*_q$. As is evident graphically and from (12), z^* is the intersection point of $\phi(\cdot)$ and the 45-degree line (the pivotal voter's indifference curve) on $[-1,1]$, provided that they have one. Because $\phi(\cdot)$ is non-decreasing and $\phi^* = \phi(v^*_q)$ by definition (cf. (5)), v^*_q and z^* must lie on different sides of ϕ^* whenever $\phi^* \in [-1,1]$ and $\phi^* \neq v^*_q$.

In what follows we assume $v^*_q, \phi^* \in [-1,1]$ to simplify exposition.[25] Let $\{a_n\}_{n\geq N}$ and $\{b_n\}_{n\geq N}$ denote the sequences of the cutoff points of the optimal censorship policies. Define $a^* := \lim_{n\to\infty} a_n$ and $b^* := \lim_{n\to\infty} b_n$. Theorem 2 shows that both limits exist and explicitly characterize them.

[25] Any $\phi^* < -1$ (resp. $\phi^* > 1$) is equivalent to the case $\phi^* = -1$ (resp. $\phi^* = 1$). The same applies for v^*_q.

Figure 5: The relationship between v_q^*, ϕ^* and z^*, and the asymptotically optimal censorship policy

(a) $\rho > 0$ and $\underline{\phi} \leq \phi^* < v_q^*$ (b) $\rho > 0$ and $\phi^* < \underline{\phi} < v_q^*$



Theorem 2. *Suppose that the single-crossing property holds and $v_q^*, \phi^* \in [-1,1]$. Define*

$$\overline{\phi} := \sup\left\{y \in [-1,1] : \mathbb{E}_F\left[k|k \leq y\right] \leq v_q^*\right\} \tag{13}$$

$$\underline{\phi} := \inf\left\{y \in [-1,1] : \mathbb{E}_F\left[k|k \geq y\right] \geq v_q^*\right\} \tag{14}$$

such that $\underline{\phi} \leq v_q^* \leq \overline{\phi}$ always holds. Then a^*, b^* and W^* are characterized as follows:

1. *If $v_q^* = \phi^* = z^*$, then $a^* = b^* = \phi^*$.*
2. *If $v_q^* > \phi^* \geq z^*$, then $a^* = z^*$ and $b^* = \max\{\phi^*, \underline{\phi}\} \leq v_q^*$.*
3. *If $v_q^* < \phi^* \leq z^*$, then $a^* = \min\left\{\phi^*, \overline{\phi}\right\} \geq v_q^*$ and $b^* = z^*$.*

The boundaries $\overline{\phi}$ and $\underline{\phi}$ defined in this theorem jointly determine the limit extent to which a monopoly sender can manipulate outcomes of large elections. More specifically, notice that under any censorship policy the induced election outcome in the limit can be characterized by an implementation threshold $t^* \in [-1,1]$ such that reform (status quo) is implemented almost surely if $k > (<) t^*$. To implement any interior $t^* \in (-1,1)$ it is necessary that $a^* \leq t^* \leq b^*$. Given the pivotal voter's type v_q^*, the set of feasible implementation thresholds in the limit is exactly $[\underline{\phi}, \overline{\phi}]$.

By Theorem 2, a^* and b^* must satisfy the following chain of inequalities:

$$\min\left\{v_q^*, z^*\right\} \leq a^* \leq \min\left\{\phi^*, z^*\right\} \leq \max\left\{\phi^*, z^*\right\} \leq b^* \leq \max\left\{v_q^*, z^*\right\}\ . \tag{15}$$

(15) reveals that both z^* and ϕ^* must lie in $[a^*, b^*]$, the limiting revelation interval as $n \to \infty$. $z^* \in [a^*, b^*]$ follows from the fact that $a_n \leq z_n \leq b_n$ for all $n \geq N$ (cf. Theorem 1) and $z^* = \lim_{n\to\infty} z_n$. The reason for $\phi^* \in [a^*, b^*]$ is the following. If $k > \phi^*$ (resp. $k < \phi^*$) then as $n \to \infty$ the sender almost surely prefers Reform (resp. Status Quo) and hence would like to induce a higher (resp. lower) posterior expected state. This cannot be efficiently achieved if any $k \neq k'$ with $k \leq \phi^* \leq k'$

are not fully separated ex-post.

Three implications are immediate from (15). First, regardless of the sender's preference, full disclosure (i.e., $a_n = -1$ and $b_n = 1$) is generically suboptimal for sufficiently large n if either (i) $v_q^* \in (-1,1)$ so that the pivotal voter's ex-ante preference is state-dependent, or (ii) $z^* \in (-1,1)$ so that the sender is not uniformly biased towards any alternative relative to the pivotal voter as $n \to \infty$. Second, no disclosure (i.e., $a_n = b_n \in \{-1,1\}$) is never optimal for sufficiently large n whenever $\phi^* \in (-1,1)$, that is, the sender's ex-ante preference is state-dependent. Third, if both $z^*, v_q^* \in (-1,1)$ hold, then $-1 < a_n < b_n < 1$ can be ensured – so that the optimal policy will indeed censor both sufficiently high and low state realizations – for n large enough. Condition $z^* \in (-1,1)$ requires that the indifference curves of the sender and pivotal voter intersect at some interior point in $(-1,1)$ as $n \to \infty$. This is the case, for instance, if the sender is self interested with $\rho = 0$ and $\chi \in (-1,1)$ (cf. Example 1), or if the sender is a Utilitarian planner (cf. Example 2) and (8) admits an interior solution on $(-1,1)$. We summarize these observations in Corollary 1.

Corollary 1. *Suppose that the single-crossing property holds. If either $v_q^* \in (-1,1)$ or $z^* \in (-1,1)$, then full disclosure is suboptimal for sufficiently large n. If $\phi^* \in (-1,1)$, then no disclosure is suboptimal for sufficiently large n. If both $z^*, v_q^* \in (-1,1)$, then the optimal information policy will censor both upwards and downwards for sufficiently large n.*

As we explained in the previous section, a necessary condition for the optimal censorship policy to have a non-trivial revelation interval (i.e., $a_n < b_n$) with finite n is that the sender is uncertain about the pivotal voter's type. As $n \to \infty$ such uncertainty vanishes and one might expect $a^* = b^*$ in some circumstances. Indeed, for the case $v_q^* = \phi^* \in (-1,1)$ – i.e., there is no ex-ante conflict of interests between the sender and the pivotal voter – Theorem 2 immediately implies $a^* = b^* = \phi^*$. It is perhaps surprising that even in the absence of any ex-ante conflict of interests the asymptotically optimal information policy is not full disclosure, but instead a binary cutoff policy that only reveals whether the state realization is above, equal or below ϕ^*, the ex-ante threshold of acceptance of the sender. Such a binary cutoff policy is indeed asymptotically optimal because it does induce the pivotal voter to implement the reform (status quo) almost surely as $n \to \infty$ whenever $k > (<)\phi^*$, which perfectly coincides with the sender's favored outcome ex-ante.[26]

Now suppose $\phi^* \neq v_q^*$ so that the ex-ante preferences of the sender and the pivotal voter are not aligned. Corollary 2 gives sufficient and necessary conditions for $a^* = b^*$ to hold in this case.

[26] Kamenica and Gentzkow (2011) made a similar observation that aligning the interests between the sender and the receiver does not necessarily imply more information disclosure by the sender. Specifically, they wrote that "*The impact of alignment on the amount of information communicated in equilibrium is also ambiguous. On the one hand, the more Receiver responds to information in a way consistent with what Sender would do, the more Sender benefits from providing information. On the other hand, when preferences are more aligned Sender can provide less information and still sway Receiver's action in a desirable direction. Hence, making preferences more aligned can make the optimal signal either more or less informative*." (pp. 2604-2605).

Corollary 2. *Suppose that the single-crossing property holds, $v_q^* \in (-1,1)$ and $v_q^* \neq \phi^*$. If $\rho = 0$, then $a^* = b^*$ if and only if $\chi \in [\underline{\phi}, \overline{\phi}]$. If $\rho > 0$ then $a^* < b^*$.*

Consider the case $\rho = 0$ first so that the sender is self interested with $\phi^* = \chi$; he prefers the reform (status quo) to be implemented in all states $k > (<)\chi$. For this case Corollary 2 claims that $a^* = b^* = \chi$ if and only if χ lies in the feasible set of implementation thresholds $[\underline{\phi}, \overline{\phi}]$. This is because by simply revealing whether k is above or below χ the sender can already sway the pivotal voter's decision in his preferred direction in all states as $n \to \infty$. Any further expansion of the revelation interval is weakly harmful in the limit as it can only make the pivotal voter choose the sender's less preferred outcome. If instead χ lies outside $[\underline{\phi}, \overline{\phi}]$, then $a^* < b^*$ so that the optimal censorship policy contains a non-trivial revelation interval even in the limit. For instance, if $\chi < \underline{\phi}$ we will have $a^* = \chi$ and $b^* = \underline{\phi} > \chi$ by Theorem 2. The expansion of b^* from χ to $\underline{\phi}$ is driven by the demand to effectively persuade the pivotal voter. The resulting structure is very similar to the judge example in Kamenica and Gentzkow (2011) in that the pooling message "$k > \underline{\phi}$" just makes the pivotal voter indifferent ex-ante. It is the difficulty to persuade voters that produces the non-trivial revelation interval in the asymptotically optimal censorship policy.

If the sender is prosocial with $\rho > 0$ (while still supposing $v_q^* \neq \phi^*$), then by Lemma 6 we have $\phi^* \neq z^*$ and hence $a^* < b^*$ by (15). Therefore, unlike a self-interested sender, a prosocial sender's optimal information policy will contain a non-trivial revelation interval as $n \to \infty$ even when the ex-ante conflicts of interests are arbitrarily small. Figure 5 illustrates the optimal revelation interval for a prosocial sender as $n \to \infty$ when $\phi^* < v_q^*$. In contrast to the previous case with $\rho = 0$, here the emergence of a non-trivial revelation interval in the limit is not driven by the difficulty of persuading the pivotal voter. It instead stems from the fact that the sender is uncertain about his preference (as it depends on voters' private types) and he must infer this through the pivotal voter's choice.

Finally, we study the sender's payoff. Let W^* denote the sender's payoff under the optimal information policy as $n \to \infty$.[27] We compare W^* with two important benchmarks $\overline{W}$ and W^{Full}. Here, $\overline{W}$ is the sender's payoff under his *omniscient* control – i.e., he directly observes state k and voters' type profile v and dictates the election outcome – as $n \to \infty$. W^{Full} is the sender's payoff under full information disclosure as $n \to \infty$. Proposition 1 gives the rankings of $\overline{W}$, W^* and W^{Full}.

Proposition 1. *Suppose $v_q^* \in (-1,1)$. Then $\overline{W}$, W^* and W^{Full} are ranked as follows:*

1. *If $v_q^* = \phi^*$, then $\overline{W} = W^* = W^{Full}$;*
2. *If $v_q^* \neq \phi^*$ and $\underline{\phi} \leq \phi^* \leq \overline{\phi}$, then $\overline{W} = W^* > W^{Full}$;*
3. *If $\phi^* < \underline{\phi}$ or $\phi^* > \overline{\phi}$, then $\overline{W} > W^* > W^{Full}$.*

To understand Proposition 1, consider first the asymptotic payoff under full information revelation, W^{Full}. Because a sender can always opt for full disclosure, $W^* \geq W^{\text{Full}}$ necessarily holds.

[27] That is, $W^* := \lim_{n \to \infty} W_n$ and W_n is the value of persuasion problem (MP) with electorate size n.

Under full disclosure the pivotal voter of type v_q^* implements the reform if $k > v_q^*$ and does not do so otherwise. Hence, whenever an ex-ante conflict of interests exists (i.e. $v_q^* \neq \phi^*$), the sender earns strictly less under full disclosure than under his optimal information policy (i.e., $W^* > W^{\text{Full}}$). Intuitively, full disclosure implies that the conflicting states between ϕ^* and v_q^* are fully revealed, while the optimal information policy should avoid doing so as much as possible.

Next, we relate W^* to the 'omniscient control' benchmark $\overline{W}$. Recall that $[\underline{\phi}, \overline{\phi}]$ is the set of feasible implementation thresholds t^* in the limit; that is, the sender can ensure reform (status quo) being elected with probability one for $k > (<) t^*$. Whenever ϕ^* falls within this set, he can secure his preferred outcome with probability one as $n \to \infty$ by setting $t^* = \phi^*$. In that case he receives payoff $\overline{W}$, which is what he could secure in the limit under omniscient control. Note that $\phi^* \in [\underline{\phi}, \overline{\phi}]$ if either (i) ϕ^* is close to v_q^* such that the ex-ante conflict of interests between the sender and the pivotal voter is low, or (ii) v_q^* is close to $\mathbb{E}_F[k]$ so that, a priori, the pivotal voter is almost indifferent between the reform and the status quo. In the latter case very weak evidence is already sufficient to persuade the pivotal voter (and thus the feasible set is large). In case ϕ^* lies below the feasible set ($\phi^* < \underline{\phi}$), the sender's preferred alternative will (as $n \to \infty$) almost surely not be elected when $k \in (\phi^*, \underline{\phi})$. The sender therefore gets strictly less than $\overline{W}$. A similar intuition applies when $\phi^* > \overline{\phi}$.

6.2 Comparative statics

In this section we derive comparative statics for how the optimal thresholds a_n and b_n vary with the sender's preference and the voting rule. Omitted proofs for this subsection are in Appendix E.

We first study the effects of shifting the sender's preference towards the reform, holding ρ fixed. Such a shift can occur, for instance, if $\rho < 1$ and χ decreases. In that case the sender's personal payoffs from reform increase and his threshold for accepting reform decreases. Proposition 2 shows how such a preference shift towards reform affects the sender's optimal censorship policy.

Proposition 2. *Suppose $\rho < 1$ and either condition (i) or (ii) in Lemma 3 holds.*[28] *Consider any $\chi_I > \chi_{II}$. Then, for sufficiently large n, as χ decreases from χ_I to χ_{II} the following holds:*

1. *a_n weakly decreases, strictly so if $a_n \in (-1, 1)$ under $\chi = \chi_I$.*
2. *b_n weakly decreases, strictly so if $b_n \in (-1, 1)$ under $\chi = \chi_I$.*

In words, as the sender's personal payoff from reform increases, his optimal censorship policy will censor fewer states downwards but more states upwards.

The intuition of Proposition 2 is as follows. On the one hand, as the sender's personal payoff from reform increases, he becomes more tempted to persuade voters to pass the reform. Therefore, b_n decreases because the sender is now tempted to manipulate voters' beliefs upwards in some states

[28] Recall that these conditions are (i) ρ is sufficiently close to 0, or (ii) both G and $1 - G$ are strictly log-concave.

that he would previously have been willing to reveal truthfully. On the other hand, such a preference shift makes the sender less tempted to persuade voters to maintain the status quo. Consequently, a_n also decreases because the sender is now willing to truthfully reveal some states he would previously censor to manipulate voters' beliefs downwards.

For a prosocial sender with $\rho > 0$, a similar preference shift towards the reform could also occur if his welfare weighting function $w(\cdot)$ decreases in the sense of first order stochastic dominance.[29] In this way the sender systematically puts more weights on voters whose ex-post type realizations are lower and hence receive higher payoffs under reform. Such a change also makes the sender favor reform more and thus be more tempted to persuade voters to pass the reform. Following the intuition discussed above, one may expect that the result in Proposition 2 continues to hold in this case. Proposition E.1 in Appendix E confirms that, under some mild conditions, this is indeed true.

Next we turn to the effects on a_n and b_n of changing voting rule q, the required vote share to pass the reform. The results depend critically on whether the sender is self-interested ($\rho = 0$) or prosocial ($\rho > 0$). We consider the self-interested case first.

Proposition 3. *Suppose $\rho = 0$ and consider any $q_I, q_{II} \in (0,1)$ with $q_{II} > q_I$. Then, for sufficiently large n, as q rises from q_I to q_{II} the following holds:*

1. a_n weakly increases, strictly so if $a_n \in (-1,1)$ under $q = q_I$.

2. b_n weakly increases, strictly so if $b_n \in (-1,1)$ under $q = q_I$.

In words, if the sender is self-interested, then increasing the required vote share to pass the reform makes him censor more states downwards but fewer states upwards.

Proposition 3 is driven by a *stringency effect*: as q increases it becomes harder to persuade the pivotal voter to pass the reform while easier to persuade her to maintain the status quo (Alonso and Câmara, 2016*a*). This is because the pivotal voter's threshold of acceptance $v^{(nq+1)}$ increases in q. For a self-interested sender who does not care about voter welfare, his best response would be to shift up both thresholds a_n and b_n. By raising b_n the sender makes the upward pooling message "$k > b_n$" more effective in persuading the pivotal voter – who is now harder to convince – to pass the reform. At the same time, the demand for the effectiveness of the downward pooling message "$k < a_n$" is lower because it is now easier to convince the pivotal voter to maintain the status quo. Therefore, by increasing a_n the sender can expand the set of states in which he can successfully persuade the pivotal voter to maintain the status quo at minor costs of reduced effectiveness.

An important implication of Proposition 3 is that, under the stringency effect alone, both a_n and b_n increase monotonically in q for sufficiently large n. Such unambiguous effects are no longer obtained once the sender cares about voter welfare. In fact, as Proposition 4 shows, the comparative statics can then go either way.

[29] More precisely, Proposition B.2 in Appendix B shows that whenever $\rho > 0$ the sender's indifference curve $\phi_n(\cdot)$ systematically shifts downwards as $w(\cdot)$ decreases in the sense of first order stochastic dominance.

Proposition 4. *Suppose $\rho > 0$, both G and $1 - G$ are strictly log-concave, n is sufficiently large, and $-1 < a_n < b_n < 1$ holds under $q = q_I$. Then there exist $q_I, q_{II} \in (0,1)$ with $q_{II} > q_I$ such that, as q increases from q_I to q_{II}, any one of the following may happen:*

1. *a_n strictly decreases and b_n strictly increases.*
2. *a_n strictly increases and b_n strictly decreases.*
3. *both a_n and b_n strictly decrease.*
4. *both a_n and b_n strictly increase.*

In words, if the sender is prosocial, then raising the required vote share to pass the reform may make him reveal more states both upwards and downwards, censor more states in both directions, or reveal more states in one direction and censor more states in the other direction.

Proposition 4 shows that for a prosocial sender an increase in the vote share required to pass the reform may shift a_n and b_n both downwards or in opposite directions. As noted above, neither case is possible under the stringency effect alone. This result is thus driven by an additional effect. For a prosocial sender, an increase in q also affects thresholds a_n and b_n through a novel *sender-preference effect*; increasing the vote share required to pass the reform induces a shift of a prosocial sender's preference towards the reform.[30] Therefore, following the intuition discussed above for the effects of shifts in the sender's preference, this induced sender-preference effect *per se* drives both a_n and b_n downwards as q increases. As a consequence, the net effect of an increase in q depends on the relative strengths of the stringency and sender preference effects. For instance, if the sender-preference effect dominates in driving a_n while the stringency effect dominates in driving b_n, then the net effect would be a strict expansion of the revelation interval $[a_n, b_n]$ on both sides.[31] Conversely, if the sender-preference effect dominates in driving b_n while the stringency effect dominates in driving a_n, then the net effect would be a strict reduction of the revelation interval $[a_n, b_n]$ on both sides.

The sender-preference effect stems from the inference problem based on the pivotal voter's choice. This can be seen as follows. Let q increase from q' to q''. The pivotal voter's type then shifts from $v^{(nq'+1)}$ to $v^{(nq''+1)}$. Under cutoff q'' the pivotal event $v^{(nq''+1)} = x$ necessarily implies $v^{(nq'+1)} \leq x$ (i.e., the pivotal voter's type must be lower than x for cutoff q'). Therefore, for any fixed x, the event that the pivotal voter's type equals x implies that the entire realized type profile is systematically lower (and thus voters' ex-post payoffs from reform becomes systematically higher) as q increases. This makes any prosocial sender leaning more towards reform.

It is important to note that the sender-preference effect exists if and only if the sender is both prosocial and imperfectly informed about voters' preferences. This effect is therefore absent in

[30] Formally, an increase in q systematically shifts the sender's indifference curve $\phi_n(\cdot)$ downwards whenever $\rho > 0$. See Proposition B.2 in Appendix B.

[31] This case can be illustrated by the shift from panel (a) to (b) in Figure 5. There, an increase in q raises both v_q^* and $\underline{\phi}$, and lowers z^* through shifting down $\phi(\cdot)$. This strictly expands the (limiting) revelation interval in both directions.

Alonso and Câmara (2016*a*), who study effects of voting rules on optimal persuasion strategies in a model where the sender is perfectly informed of voters' preferences.

7 Competition in persuasion with multiple senders

In this section we extend our model to allow for competition in persuasion with multiple senders. We show that our main result for monopolistic persuasion – that the single-crossing property ensures the optimality of censorship policies in sufficiently large elections – continues to hold under competition in persuasion. As an application, we use our results to study the welfare impact of media competition.

Specifically, we consider a setup with multiple senders competing in persuading voters a la Gentzkow and Kamenica (2017*b*). Let there be a set M of senders with $|M| \geq 2$. For each sender $m \in M$, his preference is characterized by utility function (1) with parameters $\rho_m \in [0,1]$, $\chi_m \in \mathbb{R}$, and weighting function $w_m(\cdot)$. In this way, for each sender m we can obtain his indifference curve $\phi_n^m(\cdot)$ via (3). We assume that all senders' preferences are commonly known among themselves. Each sender m simultaneously chooses an information policy π_m from the feasible set Π, prior to observing the realization of k. Given profile $\{\pi_m\}_{m\in M}$, we denote by $\pi := \left\langle \{\pi_m\}_{m\in M} \right\rangle$ the *joint information policy* induced by observing the signal realizations from all π_m's. Notice that $\pi \in \Pi$ necessarily because it is clearly feasible. In this way, our information environment is *Blackwell-connected*; given any strategy profile $\pi_{-m} := \left\{\pi_j\right\}_{j\in M\setminus\{m\}}$ of other senders, each sender m can unilaterally deviate to any feasible joint information policy that is Blackwell more informative than $\langle \pi_{-m} \rangle$.[32] We focus on equilibria in pure and weakly undominated strategies.[33] The equilibrium derivation and proofs for all results in this section are in Appendix F.

Suppose the single-crossing property holds for a sender $m \in M$ and let z_n^m and $W_n^m(\cdot)$ denote, respectively, the implied switching state and his indirect utility function.[34] Then, by Theorem 1, there exists a threshold $N_m \geq 0$ such that for all $n \geq N_m$ the monopolistically optimal information policy for sender m is a censorship policy whose revelation interval $[a_n^m, b_n^m]$ contains z_n^m. Moreover, by Lemma 5 in Section 5, for any $n \geq N_m$ there are two cutoffs $\ell_n^m \in [-1, a_n^m]$ and $r_n^m \in [b_n^m, 1]$ such

[32] Formally, for two feasible joint information policies $\pi, \pi' \in \Pi$, π is Blackwell more informative than π' if $H_\pi \succeq_{MPS} H_{\pi'}$; that is, the distribution of posterior expectations about states induced by π is a mean-preserving spread of that induced by π'.

[33] As explained in footnote 12, we focus on weakly undominated strategies to rule out a plethora of uninteresting equilibria. For competition with multiple senders, the restriction to pure strategies is often made in the literature (Gentzkow and Kamenica, 2017*a*,*b*; Mylovanov and Zapechelnyuk, 2021), but may nevertheless have substantive consequences. As noted in Gentzkow and Kamenica (2017*b*) (page 318), when senders may use mixed strategies the information environment may no longer be Blackwell-connected, which is a key property we use to characterize equilibria under competition. Li and Norman (2018) show by examples that allowing for mixed strategies indeed changes the set of equilibria.

[34] Namely, z_n^m and $W_n^m(\cdot)$ are respectively defined by (6) and (9), with $\phi_n(\cdot)$ therein replaced by $\phi_n^m(\cdot)$.

that $W_n^m(\cdot)$ is strictly convex on $[\ell_n^m, r_n^m]$ and is strictly concave elsewhere on $[-1,1]$. Theorem 3 shows that the optimality of censorship policies generalizes to competition in persuasion if all other senders use pure strategies only.

Theorem 3. *Suppose the single-crossing property holds for sender m. Then for all $n \geq N_m$, there exists a subset of censorship policies $\mathscr{P}_n^m$ such that for any pure strategy profile π_{-m} of senders other than m (which need not be censorship policies), there is a censorship policy in $\mathscr{P}_n^m$ that is sender m's best response to π_{-m}. Moreover, the best response set $\mathscr{P}_n^m$ is characterized by*

$$\mathscr{P}_n^m := \{\mathscr{P}(c,d) : \; [a_n^m, b_n^m] \subseteq [c,d] \subseteq [\ell_n^m, r_n^m]\} \tag{BR}$$

In words, $\mathscr{P}_n^m$ is the set of all censorship policies whose revelation interval (i) contains sender m's optimal revelation interval under monopolistic persuasion, and (ii) is contained by the maximal interval on which the sender's indirect utility function $W_n^m(\cdot)$ is strictly convex.

For the remainder of this section we impose the following assumption:

Assumption 1. *The following conditions hold:*

1. *The single-crossing property holds for each sender $m \in M$.*
2. *For all $m \in M$ and $n \geq 0$, $\phi_n^{m\prime}(x) < 2$ holds on $[-1,1]$.*

By Lemma 3, both conditions in Assumption 1 hold – independently of the senders' preferences and the voting rule – if both G and $1-G$ are strictly log-concave. Following Gentzkow and Kamenica (2017*b*), we say an equilibrium is *minimally informative* if the joint information policy it induces is no more Blackwell informative than any information policy that can be induced by some other equilibrium. Theorem 4 characterizes the (essentially unique) outcome induced by the minimally informative equilibria under competition in persuasion.

Theorem 4. *Suppose Assumption 1 holds and let $N := \max_{m \in M} N_m$. Then, for all $n \geq N$, the following properties hold:*

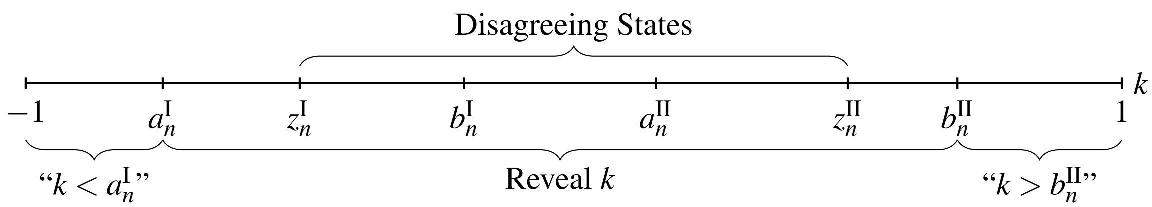
1. *In any minimally informative equilibrium the joint information policy induced by all senders is outcome equivalent to a censorship policy with revelation interval $\left[a_n^{\min}, b_n^{\max}\right]$, where $a_n^{\min} = \min_{m \in M}\{a_n^m\}$ and $b_n^{\max} = \max_{m \in M}\{b_n^m\}$.*

2. *If each sender $m \in M$ is restricted to use censorship policies from his best-response set $\mathscr{P}_n^m$ given by* (BR)*, then the minimal informative equilibrium is the unique equilibrium in pure and weakly undominated strategies.*[35]

[35] In an earlier version of this paper we establish another result: if each sender is only restricted to use censorship policies, then the minimally informative equilibrium is the unique pure strategy equilibrium that survives (two rounds of) iterated eliminations of weakly dominated strategies.

Theorem 4 implies that, under a mild regularity condition (i.e., part (2) of Assumption 1), if the single-crossing property holds for all senders then for sufficiently large elections the joint information policy induced in the minimally informative equilibrium is outcome equivalent to a censorship policy whose revelation interval is simply the convex hull of the revelation intervals of all senders' monopolistically optimal censorship policies. Figure 6 illustrates this for the case with two senders.

Figure 6: Minimally Informative Equilibrium with Two Competing Senders



Note: a_n^m and b_n^m are cutoffs of the optimal censorship policy under monopolistic persuasion for sender $m \in \{\mathrm{I}, \mathrm{II}\}$.

It is well known that multiple equilibria exist under competition in public Bayesian persuasion. The literature typically focuses on minimally informative equilibria because these are Pareto-optimal for all senders (Gentzkow and Kamenica, 2017*b*).[36] Building on part (2) of Theorem 4, we provide a novel argument for selecting the minimal informative equilibria outcome by restricting senders to use censorship policies from their best response sets. Under these restrictions, the minimally informative equilibrium is the unique equilibrium in pure and weakly undominated strategies. Due to this favorable equilibrium selection, all senders would indeed prefer these restrictions to be enforced. This would help them to avoid the risk of coordinating on equilibrium outcomes that are excessively informative and thereby would make all senders strictly worse off.

Finally, we discuss an interesting implication of Theorem 4. To do so we introduce the notion of '*disagreeing states*'. State k is a disagreeing state if there exist at least two senders $\mathrm{I}, \mathrm{II} \in M$ who are weakly biased towards different alternatives relative to the pivotal voter (formally, there exist senders $\mathrm{I}, \mathrm{II} \in M$ with $\phi_n^{\mathrm{I}}(k) \leq k \leq \phi_n^{\mathrm{II}}(k)$). These two senders thus have incentives to manipulate voters' beliefs in opposite directions. Disclosing more information then always benefits at least one of these senders.[37] This in the end leads to full revelation of all such states. When the single-crossing property holds for all senders $m \in M$, the set of disagreeing states is precisely the interval $[z_n^{\min}, z_n^{\max}]$, where $z_n^{\min} = \min_{m \in M}\{z_n^m\}$ and $z_n^{\max} = \max_{m \in M}\{z_n^m\}$. Because $a_n^{\min} \leq z_n^{\min}$ and $b^{\max} \geq z_n^{\max}$, Theorem 4 implies that all disagreeing states must be revealed in any equilibrium.

[36] An exception is Mylovanov and Zapechelnyuk (2021), who propose an equilibrium refinement based on a vanishing (entropy-based) cost of information disclosure.

[37] More precisely, as we show in Appendix F (Lemma F.2 therein), when condition (2) of Assumption 1 holds then in any disagreeing state k there exists at least one sender $m \in M$ whose indirect utility function is strictly convex at k locally. This local convexity implies positive gains from revealing more information (because it induces a mean-preserving spread on the distribution of the posterior expected state).

The implication above yields the following corollary, which gives a neat sufficient condition for full information disclosure to be the unique equilibrium outcome.

Corollary 3. *Suppose Assumption 1 holds. If there exist two senders $I, II \in M$ with $z_n^I = -1$ and $z_n^{II} = 1$, then full disclosure is the unique equilibrium outcome.*

This corollary says that full information disclosure is the unique equilibrium outcome whenever there are two senders who are uniformly biased towards different alternatives than the pivotal voter. This condition holds, for example, in the zero-sum game where competition is between two self-interested senders who always favor opposite alternatives (i.e., $\rho_{\text{I}} = \rho_{\text{II}} = 0$, $\chi_{\text{I}} \leq -1$ and $\chi_{\text{II}} \geq 1$). Such extreme conflicts of interests are, however, far from necessary to obtain full disclosure in equilibrium. This is also obtained, for example, when competition is between a 'pro-Reform' planner and an 'anti-Reform' planner under simple majority rule (cf. Examples 3 and 4 in Section 5). In this case the conflict of interests between senders is much weaker than in the previous example with opposite-minded self-interested senders; here, both planners aim at maximizing voters' payoffs and they just differ in their welfare weights.

7.1 Application: Media competition and voter welfare in large elections

Building on Corollary 3 and the asymptotic results derived in Section 6.1, we present a straightforward but insightful application to study the welfare impact of media competition.

Let $M = \{\text{I}, \text{II}\}$ and interpret these two senders as public mass media outlets. Suppose that both outlets are partisan and opposite-minded; that is, their preferred alternatives are opposite and state-independent. We model this by assuming $\rho_{\text{I}} = \rho_{\text{II}} = 0$, $\chi_{\text{I}} \leq -1$ and $\chi_{\text{II}} \geq 1$. It follows immediately from Corollary 3 that competition between these two mass media outlets will lead to full information disclosure in any equilibrium.

Now we turn to the implications for voter welfare. We focus on the limiting case $n \to \infty$. Suppose a social planner with $\rho = 1$ wants to maximize voters' utilitarian welfare. The ex-ante threshold of acceptance is thus $\phi^* = \mathbb{E}_G[v]$, the expected type of the average voter.[38] Because full disclosure is the unique equilibrium outcome, the asymptotic welfare under competition equals W^{Full}. Under the second-best benchmark (i.e., the planner can implement his own optimal information policy), the asymptotic welfare is given by W^*. By Proposition 1, $W^* \geq W^{\text{Full}}$ and a strict inequality holds whenever $v_q^* \neq \phi^*$. In other words, unless the preferences of the average and the pivotal voters are ex-ante aligned, media competition fails to maximize voters' utilitarian welfare in large elections even if it induces full information disclosure. Because full disclosure essentially makes the pivotal voter

[38] To see this, recall that the weighting function is $w(x) = x$ for $x \in [0,1]$ for a utilitarian planner. Together with $\rho = 1$, this implies $\phi^* = \int_0^1 G^{-1}(y)dw(y) = \int_{\underline{v}}^{\overline{v}} x dG(x) = \mathbb{E}_G[v]$. All analyses here apply similarly to any non-utilitarian social planner with a different weighting function $w(\cdot)$ (because this affects welfare only through ϕ^*).

the *de factor* dictator, her choice goes against the interest of the average voter in any environment where their interests are not aligned. In fact, the welfare gap $W^* - W^{\text{Full}}$ can be substantial for large differences between ϕ^* and v_q^*, which in turn depends on voting rule q (recall that v_q^* is strictly increasing while ϕ^* is invariant in q). This suggests that any welfare evaluation must take both the electoral background and institutional factors – such as the distribution of voters' preferences and voting rules – into account because they determine the ex-ante interest misalignment between the pivotal voter and the social planner. As such misalignment of interests increases, the competition between opposite-minded partisan media outlets may result in greater welfare losses due to the excessive information revelation it induces in equilibrium.

Despite its straightforwardness, our result arguably contributes to the debate on the effect of media competition on voter welfare, which is an important topic in the literature on the political economics of mass media. Most papers in this literature assume from the outset that more information is better for welfare and focus on whether media competition can improve information revelation.[39] This reasoning would lead one to conclude that media competition is ideal from a welfare perspective if it induces full information revelation. Our result suggests that this is in general not true and a careful welfare evaluation should take electoral and institutional factors into account.

8 Conclusion

This paper studies public persuasion in elections with binary alternatives. In our model, one or multiple senders can try to influence the election outcome by strategically providing public information to voters about a payoff-relevant state. Compared to prior works, our setup allows for a wide class of senders preferences and we characterize, in a unified framework, the equilibrium information provision under both monopolistic and competitive persuasion. Our main result identifies a sufficient condition that ensures the optimality of censorship policies, which reveal intermediate state realizations but censor extreme ones. This sufficient condition can be intuitively interpreted as a single-crossing property over the sender's and the pivotal voter's indifference curves. This condition holds for a sender if either the sender is self-interested, or the distribution of voters' preferences satisfies a mild regularity condition.

[39] For example, it has been argued that media competition can benefit voters and improve political accountability by increasing the costs of media capture (Besley and Prat, 2006) that aims at suppressing disclosure of unfavorable information to the politician. Competition may also discipline media outlets to provide information that aligns better with the interests of their audiences (Gentzkow and Shapiro, 2006; Chan and Suen, 2008). These papers conclude that media competition is welfare-improving because it induces better information disclosure. On the other hand, the literature also identifies channels through which media competition can deteriorate voter welfare by reducing information disclosure. For instance, competition can drive profit-maximizing media to invest fewer resources in the provision of political news or topics of common interests (Chen and Suen, 2018; Cagé, 2019; Perego and Yuksel, 2022). Innocenti (2021) shows that competition of two opposite-minded partisan media outlets may also decrease the quality of information when voters have limited attention.

Under monopolistic persuasion with a single sender, we show that censorship policy is uniquely optimal in large elections if the single-crossing property holds. Under competition in persuasion with multiple senders, the single-crossing property ensures that it is without loss of optimality for any sender to restrict attention to a subset of censorship policies, provided that all senders use pure strategies only. Moreover, under a weak regularity condition, the minimally informative pure-strategy equilibrium outcome can be reproduced by a censorship policy whose structure can be easily deduced from the monopolistically optimal censorship policies of all senders. Our analyses also produce a clean sufficient condition under which competition in persuasion can induce full information revelation as the unique equilibrium outcome.

Our results yield interesting and important normative implications. First, perhaps surprisingly, we show that full information disclosure is generically suboptimal even for a social planner who wants to maximize voters' welfare. In fact, the structure of a welfare maximizing information policy depends subtly on the planner's social preference, the electoral environment (e.g., the distributions of states and voters' preferences), and the voting rule. This observation complements the literature on media and politics by pointing out that even if media competition does induce full information revelation, this is not ideal from a welfare perspective as long as there is a conflict of interests between the average and the pivotal voters. Second, we deliver a novel insight regarding how a prosocial sender should tailor his optimal public information policy in response to changes in voting rules. We show that for a prosocial sender who is imperfectly informed about voters' preferences, increasing the required vote share for passing an alternative affects his optimal censorship policy through both a *stringency effect* – by making it more difficult to persuade the pivotal voter to pass that alternative – and a novel *sender-preference effect* – by inducing a shift of the sender's preference towards that alternative. The latter effect is absent in environments like Alonso and Câmara (2016*a*), where the sender is fully aware of voters' preferences.

We conclude by discussing some limitations of our paper and suggesting some avenues for future research. First, throughout the paper we focus on public persuasion and assume that voters' private types are unknown to any sender. These exclude the possibilities of *targeted persuasion* (i.e., sending different information to different voters) and *eliciting voters' private information* (e.g., by offering a menu of signals for voters). It is interesting to extend our analyses to incorporate either or both possibilities.[40] This would shed light on the strategic values of micro-targeting and screening for a monopoly sender.

Second, we assume that the welfare weights senders assign to each voter depend on the voter's ex-post payoff ranking, but not on her identity or any other characteristic. Under our assumption

[40] Heese and Lauermann (2021) study in a binary-state model targeted persuasion by a monopoly sender whose preference is independent of voters' private types. They show that the sender can ensure his preferred alternative to win with probability one in equilibrium as the electorate size goes to infinity. It is unclear, however, whether their results continue to hold if the sender's preference can depend on voters' private types as in our model.

that voters are ex-ante homogeneous this restriction is without loss of generality. In practice, however, voters do differ in characteristics that might be observable to senders; e.g., gender, age, region of residence, occupation, ethnic group, party affiliation, etc. These characteristics often systematically influence how voters fare under policy reforms; e.g., drivers are arguably more likely to experience greater negative income shocks from a car fuel levy. In these scenarios it is interesting to study the optimal information policy when senders' welfare weights can depend on such observable characteristics. One possible way to do so is to enrich our model by allowing voters to be heterogeneous in both observable and unobservable dimensions.

Third, in our model the policy reform affects voters homogeneously by shifting their payoffs under reform by a same amount. We impose this assumption for tractability, because in this way our information design problem can be solved using established linear programming techniques. In many real-life situations, however, policy reforms affect voters in heterogeneous, and sometimes even opposite, ways. For instance, citizens living in more polluted areas may benefit more from an environmental protection policy. As an alternative example, in the Brexit referendum some citizens might prefer a harder Brexit while others might instead desire a softer one. In these scenarios policy reforms can affect the distribution of voters' payoffs, and sometimes may even induce preference changes in opposite directions among citizens and thus lead to polarized attitudes towards reforms. All these features are important issues in discourses of contemporary distributive politics. We therefore believe that exploring the implications of heterogeneous policy effects for information design in elections is a promising and highly relevant area for future research.

We hope that the theoretical framework and results of our paper can serve as a good starting point to explore these additional research questions.

References

Alonso, Ricardo and Odilon Câmara. 2016*a*. "Persuading voters." *American Economic Review* 106(11):3590–3605.

Alonso, Ricardo and Odilon Câmara. 2016*b*. "Political disagreement and information in elections." *Games and Economic Behavior* 100:390–412.

Arieli, Itai, Yakov Babichenko, Rann Smorodinsky and Takuro Yamashita. 2022. "Optimal Persuasion via Bi-pooling." forthcoming in Theoretical Economics.

Ashworth, Scott and Ethan Bueno De Mesquita. 2009. "Elections with platform and valence competition." 67(1):191–216.

Au, Pak Hung and Keiichi Kawai. 2020. "Competitive information disclosure by multiple senders." *Games and Economic Behavior* 119:56–78.

Au, Pak Hung and Keiichi Kawai. 2021. "Competitive disclosure of correlated information." *Economic Theory* 72(3):767–799.

Bagnoli, Mark and Ted Bergstrom. 2005. "Log-concave probability and its applications." *Economic theory* 26(2):445–469.

Bardhi, Arjada and Yingni Guo. 2018. "Modes of persuasion toward unanimous consent." *Theoretical Economics* 13(3):1111–1149.

Bergemann, Dirk and Stephen Morris. 2016. "Bayes correlated equilibrium and the comparison of information structures in games." *Theoretical Economics* 11(2):487–522.

Bergemann, Dirk and Stephen Morris. 2019. "Information design: A unified perspective." *Journal of Economic Literature* 57(1):44–95.

Besley, Timothy and Andrea Prat. 2006. "Handcuffs for the grabbing hand? The role of the media in political accountability." *American Economic Review* 96(3):720–736.

Cagé, Julia. 2019. "Media competition, information provision and political participation: Evidence from French local newspapers and elections, 1944–2014." *Journal of Public Economics* 185:104077.

Chakraborty, Archishman and Parikshit Ghosh. 2016. "Character endorsements and electoral competition." 8(2):277–310.

Chakraborty, Archishman, Parikshit Ghosh and Jaideep Roy. 2020. "Expert-Captured Democracies." 110(6):1713–51.

Chan, Jimmy, Seher Gupta, Fei Li and Yun Wang. 2019. "Pivotal persuasion." *Journal of Economic theory* 180:178–202.

Chan, Jimmy and Wing Suen. 2008. "A spatial theory of news consumption and electoral competition." *Review of Economic Studies* 75(3):699–728.

Chen, Heng and Wing Suen. 2018. "Competition for Attention in the News Media Market." Working paper, University of Hong Kong.

Cui, Zhihan and Dilip Ravindran. 2020. "Competing persuaders in zero-sum games." Working paper. Columbia University.

Dizdar, Deniz and Eugen Kováč. 2020. "A simple proof of strong duality in the linear persuasion problem." *Games and Economic Behavior* 122:407–412.

Dworczak, Piotr and Anton Kolotilin. 2019. "The persuasion duality." Working paper.

Dworczak, Piotr and Giorgio Martini. 2019. "The simple economics of optimal persuasion." *Journal of Political Economy* 127(5):1993–2048.

Feddersen, Timothy J and Wolfgang Pesendorfer. 1996. "The Swing Voter's Curse." *American Economic Review* 86(3):408–424.

Feddersen, Timothy and Wolfgang Pesendorfer. 1997. "Voting behavior and information aggregation in elections with private information." *Econometrica* 65(5):1029–1058.

Ferguson, Sean. 2020. "How Informed Should Voters Be?" Working paper.

Gentzkow, Matthew and Emir Kamenica. 2016. "A Rothschild-Stiglitz approach to Bayesian persuasion." *American Economic Review* 106(5):597–601.

Gentzkow, Matthew and Emir Kamenica. 2017*a*. "Bayesian persuasion with multiple senders and rich signal spaces." *Games and Economic Behavior* 104:411–429.

Gentzkow, Matthew and Emir Kamenica. 2017*b*. "Competition in Persuasion." *The Review of Economic Studies* 84(1):300–322.

Gentzkow, Matthew and Jesse M Shapiro. 2006. "Media bias and reputation." *Journal of Political Economy* 114(2):280–316.

Ginzburg, Boris. 2019. "Optimal Information Censorship." *Journal of Economic Behavior and Organization* 163:377–385.

Gitmez, A Arda and Pooya Molavi. 2022. "Polarization and media bias." *arXiv preprint arXiv:2203.12698* .

Gradwohl, Ronen, Yuval Heller and Arye L. Hillman. 2022. "Social Media and Democracy." Available at SSRN: https://ssrn.com/abstract=4149316 or http://dx.doi.org/10.2139/ssrn.4149316.

Groseclose, Tim. 2001. "A model of candidate location when one candidate has a valence advantage." 45(4):862–886.

Heese, Carl and Stephan Lauermann. 2021. "Persuasion and Information Aggregation in Large Elections." Working Paper.

Innocenti, Federico. 2021. Can Media Pluralism Be Harmful to News Quality? Technical report University of Mannheim.

Ivanov, Maxim. 2020. "Optimal monotone signals in Bayesian persuasion mechanisms." *Economic Theory* pp. 1–46.

Kamenica, Emir. 2019. "Bayesian persuasion and information design." *Annual Review of Economics* 11:249–272.

Kamenica, Emir and Matthew Gentzkow. 2011. "Bayesian persuasion." *American Economic Review* 101(6):2590–2615.

Kartik, Navin and Richard Van Weelden. 2019. "Informative cheap talk in elections." *The Review of Economic Studies* 86(2):755–784.

Kerman, Toygar, P Herings and Dominik Karos. 2020. "Persuading Strategic Voters." Working paper.

Kleiner, Andreas, Benny Moldovanu and Philipp Strack. 2021. "Extreme points and majorization: Economic applications." *Econometrica* 89(4):1557–1593.

Kolotilin, Anton. 2018. "Optimal information disclosure: A linear programming approach." *Theoretical Economics* 13(2):607–635.

Kolotilin, Anton, Timofiy Mylovanov and Andriy Zapechelnyuk. 2022. "Censorship as optimal persuasion." *Theoretical Economics* 17(2):561–585.

Kolotilin, Anton, Tymofiy Mylovanov, Andriy Zapechelnyuk and Ming Li. 2017. "Persuasion of a privately informed receiver." *Econometrica* 85(6):1949–1964.

Li, Fei and Peter Norman. 2018. "On Bayesian persuasion with multiple senders." *Economics Letters* 170:66–70.

Li, Fei and Peter Norman. 2021. "Sequential persuasion." *Theoretical Economics* 16(2):639–675.

Liu, Shuo. 2019. "Voting with Public Information." *Games and Economic Behavior* 113:694–719.

Mathevet, Laurent, Jacopo Perego and Ina Taneva. 2020. "On information design in games." *Journal of Political Economy* 128(4):1370–1404.

Mylovanov, Tymofiy and Andriy Zapechelnyuk. 2021. "A model of debate: Moderation vs free speech." working paper.

Perego, Jacopo and Sevgi Yuksel. 2022. "Media competition and social disagreement." *Econometrica* 90(1):223–265.

Quiggin, John. 1982. "A theory of anticipated utility." *Journal of Economic Behavior & Organization* 3(4):323–343.

Schnakenberg, Keith E. 2015. "Expert advice to a voting body." *Journal of Economic Theory* 160:102 – 113.

Schnakenberg, Keith E. 2017. "Informational Lobbying and Legislative Voting." *American Journal of Political Science* 61:129–145.

Sun, Junze. 2022*a*. "Full information disclosure in linear persuasion games." Working paper.

Sun, Junze. 2022*b*. "On the optimality of censorship in linear persuasion games." Working paper.

Sun, Junze, Arthur Schram and Randolph Sloof. 2021. "Why Biased Endorsements Can Manipulate Elections." Available at SSRN: https://ssrn.com/abstract=3420703 or http://dx.doi.org/10.2139/ssrn.3420703.

Taneva, Ina A. 2019. "Information Design." *American Economic Journal: Microeconomics* 11(4):151–85.

Van der Straeten, Karine and Takuro Yamashita. 2020. "On the veil-of-ignorance principle: welfare-optimal information disclosure in voting." working paper.

Wang, Yun. 2013. "Bayesian persuasion with multiple receivers." Working paper. Available at SSRN: https://ssrn.com/abstract=2625399 or http://dx.doi.org/10.2139/ssrn.2625399.

Appendices (for online publication)

Contents

A Derivations and relevant properties of $\hat{G}_n(\cdot;q)$

In this appendix we formally derive $\hat{G}_n(\cdot;q)$ and establish its relevant properties in Proposition A.1, which imply Lemma 1 in Section 3. For all $y,q \in [0,1]$, define

$$\tau_n(y;q) := \frac{(n+1)!}{\lfloor x \rfloor ! \cdot \lceil n(1-q) \rceil !} y^{\lfloor nq \rfloor} (1-y)^{\lceil n(1-q) \rceil} \tag{A.1}$$

where $\lfloor \cdot \rfloor$ and $\lceil \cdot \rceil$ denote, respectively, the floor and ceiling functions. In fact, $\tau_n(\cdot;q)$ is the density function of a Beta distribution $B(\alpha,\beta)$ with parameters $\alpha = \lfloor nq \rfloor + 1$ and $\beta = \lceil n(1-q) \rceil + 1$. The following properties about $\tau_n(y;q)$ are useful.

Lemma A.1. *Suppose nq is an integer. Then the following properties hold:*

(a) $\tau_n'(y;q) = \tau_n(y;q)\frac{n(q-y)}{y(1-y)}$.

(b) $\tau_n(y;q)$ *is increasing on* $[0,q)$ *and decreasing on* $(q,1]$.

(c) $\lim_{n\to\infty} \tau_n(y;q) = \infty$ *if* $y = q$ *and* $\lim_{n\to\infty} \tau_n(y;q) = 0$ *if* $y \neq q$.

Proof of Lemma A.1. Since nq is an integer, we can drop the floor and ceiling functions in (A.1). Taking the natural logarithm of $\tau_n(y;q)$ and computing its derivative then yields

$$\frac{\tau_n'(y;q)}{\tau_n(y;q)} = \frac{n(q-y)}{y(1-y)}$$

Hence, $\tau_n'(y;q) > (<) 0$ for $y < (>) q$. This proves (a) and (b). To show (c), we use Stirling's formula to approximate $n!$ for all positive integer n: $n! \approx \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$.[1] With this approximation, we obtain

$$\tau_n(y;q) \approx \sqrt{\frac{n}{2\pi q(1-q)}} \left(\frac{y}{q}\right)^{nq} \left(\frac{1-y}{1-q}\right)^{n(1-q)} \tag{A.2}$$

If $y = q$, then $\tau_n(y;q) \approx \sqrt{\frac{n}{2\pi q(1-q)}} \to \infty$. If $y \neq q$, we take the natural logarithm of (A.2) and get

$$\ln \tau_n(y;q) \approx \frac{1}{2}\ln n + n\psi(y;q) - \frac{1}{2}\ln 2\pi q(1-q) \tag{A.3}$$

where

$$\psi(y;q) := q\ln\frac{y}{q} + (1-q)\ln\frac{1-y}{1-q} \tag{A.4}$$

It holds that (i) $\psi(q;q) = 0$, and (ii) $\psi'(y;q) > (<) 0$ for $y < (>) q$. Therefore, if $y \neq q$, then $\psi(y;q) < 0$ and the right hand side of (A.3) converges to $-\infty$ as $n \to \infty$. This implies $\lim_{n\to\infty} \tau_n(y;q) = 0$ for $y \neq q$ and thus completes the proof for part (c). □

[1] The expression $l_n \approx r_n$ denotes $\lim_{n\to\infty} \frac{l_n}{r_n} = 1$, where l_n and r_n are real number sequences.

Lemma A.1 immediately extends to other values of q in which nq is not an integer. In this case, we can just replace q by $\hat{q} := \frac{\lfloor nq \rfloor}{n}$. In this way, (a) and (b) of Lemma A.1 hold with $\hat{q}$. Part (c) of Lemma A.1 also holds for $\hat{q}$ because $\hat{q}$ converges to q as $n \to \infty$. In the remainder of this appendix and all subsequent appendices we assume nq to be an integer for ease of exposure, with the understanding that this is without loss of generality.

Now we are ready to derive $\hat{G}_n(\cdot;q)$, the distribution of the pivotal voter's type $v^{(nq+1)}$. Let $\hat{g}_n(\cdot;q)$ denote the density function. Consider $x \in [\underline{v}, \overline{v}]$. For $v^{(nq+1)} = x$ to hold, there must be nq voters with $v_i \leq x$ and $n(1-q)$ others with $v_i \geq x$, with the remaining pivotal voter having $v_i = x$. Because voters' types are independently drawn from G, we have

$$\hat{g}_n(x;q) = \frac{(n+1)!}{(nq)![n(1-q)]!}(G(x))^{nq}(1-G(x))^{n(1-q)}g(x) = \tau_n\left(G(x);q\right)g(x) \tag{A.5}$$

and

$$\hat{G}_n(x;q) = \int_{\underline{v}}^{x} \tau_n\left(G(x);q\right)g(x)dx = \int_0^{G(x)} \tau_n(y;q)dy \tag{A.6}$$

We prove the following proposition about $\hat{G}_n(x;q)$.

Proposition A.1. *Let $v_q^* := G^{-1}(q)$. The following properties hold:*

1. $\hat{G}_n(\cdot;q)$ is strictly increasing and $v^{(nq+1)}$ converges in probability to v_q^*.

2. $\hat{g}_n(\cdot;q)$ is single-peaked for all $q \in (0,1)$ when n is sufficiently large. In addition, if $g(\cdot)$ is strictly log-concave, then $\hat{g}_n(\cdot;q)$ is strictly log-concave for all $n \geq 0$ and $q \in (0,1)$.

Statement (1) of this proposition implies Lemma 1 in Section 3. Statement (2) says that regardless of the shape of G and voting rule q, for a sufficiently large electorate the distribution of the pivotal voter will be single-peaked. Moreover, large n is not needed if g is already log-concave. This property will be exploited in Appendix C.2 for the proof of Lemma 5.

Proof of Proposition A.1. We first show part (1). The fact that $\hat{G}_n(\cdot;q)$ is strictly increasing follows immediately from $\hat{g}_n(x;q) = \tau_n(G(x);q)g(x) > 0$. To show that $v^{(nq+1)}$ converges in probability to v_q^*, it suffices to establish

$$\lim_{n\to\infty} \hat{G}_n(x;q) \to \begin{cases} 0, & \text{if } x < v_q^* \\ 1/2, & \text{if } x = v_q^* \\ 1, & \text{if } x > v_q^* \end{cases} \tag{A.7}$$

For $x < v_q^*$ we have $G(x) < q$ and

$$\hat{G}_n(x;q) = \int_0^{G(x)} \tau_n(y;q)dy < G(x)\tau_n(G(x);q) \to 0$$

the second and third steps of which follow from (b) and (c) of Lemma A.1, respectively. If instead $x > v_q^*$, then $G(x) > q$ and $\int_{G(x)}^1 \tau_n(y)dy < (1-G(x))\tau_n(G(x);q) \to 0$. Therefore, $\hat{G}_n(x) = 1 - \int_{G(x)}^1 \tau_n(y)dy \to 1$. Finally, if $x = v_q^*$, then $G(x) = G(v_q^*) = q$ and $\hat{G}_n(x) = \int_0^q \tau_n(y;q)dy$. Below we show $\lim_{n\to\infty} \int_0^q \tau_n(y;q)dy = 1/2$.

Recall that $\tau_n(y;q)$ is the density function of a random variable Y following Beta distribution $B(\alpha,\beta)$ with parameters $\alpha = nq+1$ and $\beta = n(1-q)+1$. Let q_n denote the median of Y; that is, $\int_0^{q_n} \tau_n(y;q)dy = 1/2$. We show that the sequence of medians q_n converges to q and thus $\lim_{n\to\infty} \int_0^q \tau_n(y;q)dy = \lim_{n\to\infty} \int_0^{q_n} \tau_n(y;q)dy = 1/2$. For a Beta-distributed random variable $Y \sim Beta(\alpha,\beta)$, Groeneveld and Meeden (1977) show that its median q_n must be bounded between its mean μ_n and mode m_n. For a Beta distribution, it is also well known that $\mu_n = \frac{\alpha}{\alpha+\beta}$ and $m_n = \frac{\alpha-1}{\alpha+\beta-2}$. Since $\alpha = nq+1$ and $\beta = n(1-q)+1$, both μ_n and m_n converge to q as $n\to\infty$. This implies that the median q_n must converge to q as well. This establishes part (1) of this proposition.

Next we prove part (2). By (A.5) and part (a) of Lemma A.1, we have

$$\begin{aligned} \hat{g}_n'(x;q) &= \tau_n'(G(x);q)g^2(x) + \tau_n(G(x);q)g'(x) \\ &= \hat{g}_n(x;q)\left(n\frac{g(x)}{G(x)}\frac{q-G(x)}{1-G(x)} + \frac{g'(x)}{g(x)}\right) \end{aligned} \tag{A.8}$$

Therefore,

$$\frac{\hat{g}_n'(x;q)}{\hat{g}_n(x;q)} = n\frac{g(x)}{G(x)}\frac{q-G(x)}{1-G(x)} + \frac{g'(x)}{g(x)} \tag{A.9}$$

We show that $\hat{g}_n(\cdot;q)$ is single-peaked for sufficiently large n. By (A.9),

$$\hat{g}_n'(x;q) > 0 \Longleftrightarrow \lambda_n(x) := q - G(x) + \frac{1}{n}\frac{G(x)\,(1-G(x))}{g(x)}\frac{g'(x)}{g(x)} > 0$$

Recall that g is strictly positive and twice-continuously differentiable on $[\underline{v},\overline{v}]$. These imply that (i) there exists some $\varepsilon > 0$ such that $g(x) > \varepsilon$ for all x, and (ii) both $\frac{G(x)(1-G(x))}{g(x)}\frac{g'(x)}{g(x)}$ and its first order derivative are uniformly bounded. Therefore, as $n\to\infty$, $\lambda_n(x)$ and $\lambda_n'(x)$ converge uniformly to $q - G(x)$ and $-g(x)$, respectively. Hence, for sufficiently large n, $\lambda_n(x)$ is strictly decreasing and its root $\hat{x}_n$ converges to v_q^*. This implies that $\hat{g}_n(x;q)$ is single-peaked for sufficiently large n.

Finally, we assume g is strictly log-concave and show that in this case $\hat{g}_n(\cdot;q)$ is also strictly log-concave for all $n \geq 0$ and $q \in (0,1)$. Our starting point is that the strict log-concavity of g implies that both G and $1-G$ are strictly log-concave (see, for instance, Bagnoli and Bergstrom (2005)). Therefore, $\frac{g(\cdot)}{G(\cdot)}$ is strictly decreasing and by (A.9) it suffices to show that function $\eta(x) := \frac{g(x)}{G(x)}\frac{q-G(x)}{1-G(x)}$ is strictly decreasing. Differentiating $\eta(x)$ and simplifying the expression, we obtain

$$\eta'(x) = \frac{g(x)}{G(x)}\frac{1-q}{1-G(x)}\left\{\left(\frac{g(x)}{G(x)} - \frac{g'(x)}{g(x)}\right)\frac{G(x)-q}{1-q} - \frac{g(x)}{1-G(x)}\right\}$$

Therefore, $\eta'(x) < 0$ if and only if

$$\left(\frac{g(x)}{G(x)} - \frac{g'(x)}{g(x)}\right)\frac{G(x)-q}{1-q} < \frac{g(x)}{1-G(x)}$$

Recall that $\frac{g(x)}{G(x)} - \frac{g'(x)}{g(x)} > 0$ because G is strictly log-concave. The above inequality thus holds for all $q \geq G(x)$. Now we establish the above inequality for $q < G(x)$. Using the fact that $\frac{G(x)-q}{1-q}$ is strictly decreasing in q for $q < G(x)$, we get

$$\left(\frac{g(x)}{G(x)} - \frac{g'(x)}{g(x)}\right)\frac{G(x)-q}{1-q} < \left(\frac{g(x)}{G(x)} - \frac{g'(x)}{g(x)}\right)G(x) \tag{A.10}$$

It is therefore sufficient to show that the right-hand side of (A.10) is strictly smaller than $\frac{g(x)}{1-G(x)}$. This is true because

$$\begin{aligned}\left(\frac{g(x)}{G(x)} - \frac{g'(x)}{g(x)}\right)G(x) < \frac{g(x)}{1-G(x)} &\Longleftrightarrow \frac{g^2(x)-g'(x)G(x)}{g(x)} < \frac{g(x)}{1-G(x)} \\ &\Longleftrightarrow g^2(x)\left(1-G(x)\right) - g'(x)G(x)\left(1-G(x)\right) < g^2(x) \\ &\Longleftrightarrow -G(x)g^2(x) - g'(x)G(x)\left(1-G(x)\right) < 0 \\ &\Longleftrightarrow 0 < g^2(x) + g'(x)\left(1-G(x)\right) \\ &\Longleftrightarrow 1-G \text{ is strictly log-concave} \\ &\Longleftarrow g \text{ is strictly log-concave}\end{aligned}$$

This completes the proof. □

B Useful properties of $\phi_n(\cdot)$ and omitted proofs for Section 4

In this appendix we derive and establish some important properties for $\phi_n(\cdot)$ – the indifference curve of the sender – and its limit as $n \to \infty$. We also prove Lemmas 2 and 3 in Section 4.

For each $j \in \{1, \cdots, n+1\}$ and $x \in [\underline{v}, \overline{v}]$, let

$$\varphi_j(x;q,n) := \mathbb{E}\left[v^{(j)} \middle| v^{(nq+1)} = x; q, n\right]$$

denote the expectation of $v^{(j)}$ conditional on event $v^{(nq+1)} = x$. By (3) in Section 4 we have

$$\phi_n(x) := \mathbb{E}\left[\varphi_n(v) \middle| v^{(nq+1)} = x\right] = \rho \sum_{j=1}^{n+1} w_j \varphi_j(x;q,n) + (1-\rho)\chi \tag{B.1}$$

If $\rho = 0$, it is obvious that $\phi_n(x) = \chi$ is a constant. If $\rho > 0$, the properties of $\phi_n(x)$ depend closely on $\varphi_j(x;q,n)$. For any $j \neq nq+1$, let $\widetilde{g}_j(\cdot|x;q,n)$ denote the density function for the distribution of $v^{(j)}$ conditional on $v^{(nq+1)} = x$ given parameters q and n. We show that

$$\widetilde{g}_j(y|x;q,n) = \begin{cases} \tau_{nq-1}\left(\frac{G(y)}{G(x)}; \frac{j-1}{nq}\right)\frac{g(y)}{G(x)}, & \text{if } j < nq+1 \\ \tau_{n(1-q)-1}\left(\frac{G(y)-G(x)}{1-G(x)}; \frac{j-nq-2}{n(1-q)}\right)\frac{g(y)}{1-G(x)}, & \text{if } j > nq+1 \end{cases}. \tag{B.2}$$

To see why, first consider $j < nq+1$. Conditional on $v^{(nq+1)} = x$, $v^{(j)}$ is the j-th lowest order statistic from nq independent random draws from a truncated distribution with cdf $\frac{G(y)}{G(x)}$ for $y \in [\underline{v}, x]$. (B.2) for $j < nq+1$ thus follows from (A.5). Now consider $j > nq+1$. Conditional on $v^{(nq+1)} = x$, $v^{(j)}$ is the $(j-nq-1)$-th lowest order statistic from $n(1-q)$ independent random draws from a truncated distribution with cdf $\frac{G(y)-G(x)}{1-G(x)}$ for $y \in [x, \overline{v}]$. This implies (B.2) for $j < nq+1$ through (A.5). Lemma B.1 explicitly characterizes $\varphi_j(x;q,n)$.

Lemma B.1. *For all $x \in [\underline{v}, \overline{v}]$,*

$$\varphi_j(x;q,n) = \begin{cases} \int_0^1 \underline{t}(x,y)\tau_{nq-1}\left(y; \frac{j-1}{nq}\right)dy, & \text{if } j < nq+1 \\ x, & \text{if } j = nq+1 \\ \int_0^1 \overline{t}(x,y)\tau_{n(1-q)-1}\left(y; \frac{j-nq-2}{n(1-q)}\right)dy, & \text{if } j > nq+1 \end{cases} \tag{B.3}$$

where

$$\underline{t}(x,y) := G^{-1}(yG(x)) \tag{B.4}$$

$$\overline{t}(x,y) := G^{-1}(y + (1-y)G(x)) \tag{B.5}$$

for all $x \in [\underline{v}, \overline{v}]$ and $y \in [0,1]$.

Proof of Lemma B.1. $\varphi_j(x;q,n) = x$ for $j = nq+1$ follows immediately from its definition. For $j < nq+1$, it follows from (B.2) that

$$\begin{aligned} \varphi_j(x;q,n) &= \int_{\underline{v}}^x y\tilde{g}_j(y|x;q,n)dy = \int_{\underline{v}}^x y\tau_{nq-1}\left(\frac{G(y)}{G(x)}; \frac{j-1}{nq}\right)\frac{dG(y)}{G(x)} \\ &= \int_0^1 G^{-1}(yG(x))\,\tau_{nq-1}\left(y; \frac{j-1}{nq}\right)dy = \int_0^1 \underline{t}(x,y)\tau_{nq-1}\left(y; \frac{j-1}{nq}\right)dy \end{aligned}$$

Finally, for all $j > nq+1$ it follows from (B.2) that

$$\begin{aligned}\varphi_j(x;q,n) &= \int_x^{\overline{v}} y\tilde{g}_j(y|x;q,n)dy = \int_x^{\overline{v}} y\tau_{n(1-q)-1}\left(\frac{G(y)-G(x)}{1-G(x)};\frac{j-nq-2}{n(1-q)}\right)\frac{dG(y)}{1-G(x)} \\ &= \int_0^1 G^{-1}\left(y+(1-y)G(x)\right)\tau_{n(1-q)-1}\left(y;\frac{j-nq-2}{n(1-q)}\right)dy \\ &= \int_0^1 \bar{t}(x,y)\tau_{n(1-q)-1}\left(y;\frac{j-nq-2}{n(1-q)}\right)dy\end{aligned}$$

This completes the proof. □

Lemma B.2 summarizes useful properties about the functions $\underline{t}(x,y)$ and $\bar{t}(x,y)$ defined above.

Lemma B.2. *Both $\underline{t}(x,y)$ and $\bar{t}(x,y)$ are three times continuously differentiable and they satisfy the following properties:*

1. *$\underline{t}(x,y) < x < \bar{t}(x,y)$ for all $y \in (0,1)$.*
2. *Both $\underline{t}(x,y)$ and $\bar{t}(x,y)$ are strictly increasing in x and y.*
3. *If G is strictly log-concave, then $\underline{t}_x(x,y) < 1$ for all $y \in (0,1)$.*
4. *If $1-G$ is strictly log-concave, then $\bar{t}_x(x,y) < 1$ for all $y \in (0,1)$.*

Proof of Lemma B.2. The fact that both $\underline{t}(x,y)$ and $\bar{t}(x,y)$ are three times continuously differentiable for all $(x,y) \in [\underline{v},\overline{v}] \times [0,1]$ follows from our assumption that G is twice continuously differentiable on $[\underline{v},\overline{v}]$. Parts (1) and (2) of this lemma follow immediately from the definitions of $\underline{t}(x,y)$ and $\bar{t}(x,y)$. To show part (3), note from (B.4) that

$$G\left(\underline{t}(x,y)\right) = yG(x)$$

Taking the first order derivative with respect to x on both sides and rearranging terms yields

$$g\left(\underline{t}(x,y)\right)\underline{t}_x(x,y) = yg(x) \iff \underline{t}_x(x,y) = y\frac{g(x)}{g\left(\underline{t}(x,y)\right)} = \frac{g(x)}{G(x)}\Bigg/\frac{g\left(\underline{t}(x,y)\right)}{G\left(\underline{t}(x,y)\right)} \tag{B.6}$$

If G is strictly log-concave, then $\frac{g(\cdot)}{G(\cdot)}$ is strictly decreasing. Since $\underline{t}_x(x,y) < x$ for $y \in (0,1)$, it follows from (B.6) that $\underline{t}_x(x,y) < 1$. To show part (4), note from (B.5) that

$$G\left(\bar{t}(x,y)\right) = y + (1-y)\left(1-G(x)\right)$$

holds for all x and y. Simple algebra reveals that

$$\bar{t}_x(x,y) = (1-y)\frac{g(x)}{g\left(\bar{t}(x,y)\right)} = \frac{g(x)}{1-G(x)}\Bigg/\frac{g(\bar{t}(x,y))}{1-G(\bar{t}(x,y))} \tag{B.7}$$

If $1-G$ is strictly log-concave, then $\frac{g(\cdot)}{1-G(\cdot)}$ is strictly increasing. Since $\bar{t}(x,y) > x$ for $y \in (0,1)$, it follows from (B.7) that $\bar{t}_x(x,y) < 1$. □

Notice that parameters j and q affect $\varphi_j(x;q,n)$ only through their impacts on $\widetilde{g}_j(\cdot|x;q,n)$. The next lemma shows that $\widetilde{g}_j(\cdot|x;q,n)$ increases in strict monotone likelihood-ratio dominance order as j increases and q decreases. For two probability density functions $l(\cdot)$ and $r(\cdot)$, we write $l(\cdot) \succ_{LR} r(\cdot)$ if the likelihood ratio $\frac{l(\cdot)}{r(\cdot)}$ is strictly increasing.

Lemma B.3. *The following properties for $\widetilde{g}_j(\cdot|x;q,n)$ hold:*

1. *Suppose $j' > j$, then $\widetilde{g}_{j'}(\cdot|x;q,n) \succ_{LR} \widetilde{g}_j(\cdot|x;q,n)$ holds if $j > nq+1$ or $j' < nq+1$.*
2. *Suppose $q' > q$, then $\widetilde{g}_j(\cdot|x;q,n) \succ_{LR} \widetilde{g}_j(\cdot|x;q',n)$ holds if $j < nq+1$ or $j > nq'+1$.*

Proof of Lemma B.3. We first show part (1). Using (B.2) and (A.1), we obtain

$$\frac{\widetilde{g}_{j'}(y|x;q,n)}{\widetilde{g}_j(y|x;q,n)} \propto \begin{cases} \left(\frac{G(y)}{G(x)-G(y)}\right)^{j'-j} \text{ for } y \in [\underline{v}, x], & \text{if } nq+1 > j' > j \\ \left(\frac{G(y)-G(x)}{1-G(y)}\right)^{j'-j} \text{ for } y \in [x, \overline{v}], & \text{if } j' > j > nq+1 \end{cases}$$

In both cases, the likelihood ratio $\frac{\widetilde{g}_{j'}(y|x;q,n)}{\widetilde{g}_j(y|x;q,n)}$ is strictly increasing in y since $j' > j$. To show part (2), suppose $q' > q$ and note that

$$\frac{\widetilde{g}_j(y|x;q,n)}{\widetilde{g}_j(y|x;q',n)} \propto \begin{cases} \left(\frac{G(x)}{G(x)-G(y)}\right)^{n(q'-q)} \text{ for } y \in [\underline{v}, x], & \text{if } j < nq+1 \\ \left(\frac{G(y)-G(x)}{1-G(y)}\right)^{n(q'-q)} \text{ for } y \in [x, \overline{v}], & \text{if } j > nq'+1 \end{cases}$$

In both cases, the likelihood ratio $\frac{\widetilde{g}_j(y|x;q,n)}{\widetilde{g}_j(y|x;q',n)}$ is strictly increasing in y when $q' > q$. □

With these Lemmas we can establish Proposition B.1, which collects important properties of $\varphi_j(x;q,n)$.

Proposition B.1. *Let $j \in \{1, \cdots, n+1\}$. $\varphi_j(x;q,n)$ satisfies the following properties:*

1. *$\varphi_j(x;q,n)$ is strictly increasing in index j and $\varphi_j(x;q,n) = x$ for $j = nq+1$;*
2. *$\varphi_j(x;q,n)$ is strictly increasing in x and decreasing in q for all j;*
3. *If G is strictly log-concave, then $\varphi_j'(x;q,n) < 1$ for all $j < nq+1$;*
4. *If $1-G$ is strictly log-concave, then $\varphi_j'(x;q,n) < 1$ for all $j > nq+1$.*

Proof of Proposition B.1. We start with part (1). $\varphi_{nq+1}(x;q,n) = x$ follows immediately from the definition. Moreover, (B.3) and the fact that $\underline{t}(x,y) < x < \bar{t}(x,y)$ for $y \in (0,1)$ imply $\varphi_j(x;q,n) > (<) x$ for $j > (<) nq+1$. Hence, $\varphi_{j'}(x;q,n) > \varphi_j(x;q,n)$ holds for $j' \geq nq+1 \geq j$ with at least one inequality holding strictly. Now consider $j' > j > nq+1$ or $nq+1 > j' > j$. Observe that both

$\underline{t}(x,y)$ and $\bar{t}(x,y)$ are strictly increasing functions of y, and $\varphi_j(x;q,n)$ equals the expectation of $\underline{t}(x,y)$ or $\bar{t}(x,y)$ for random variable y under distribution $\tilde{g}_j(\cdot|x;q,n)$. By Lemma B.3.1, $\tilde{g}_{j'}(\cdot|x;q,n) \succ_{LR} \tilde{g}_j(\cdot|x;q,n)$ and strict likelihood ratio dominance implies $\varphi_{j'}(x;q,n) > \varphi_j(x;q,n)$ (see, for instance, Appendix B of Krishna (2009)).

To show part (2), note that both $\underline{t}(x,y)$ and $\bar{t}(x,y)$ strictly increase in x for all $y \in (0,1)$ (cf. Lemma B.2). It then follows from (B.3) that $\varphi_j(x;q,n)$ strictly increases in x. To show that $\varphi_j(x;q,n)$ decreases in q, consider two different q' and q'' with $q' < q''$. If $nq'+1 \leq j \leq nq''+1$ then by (B.3) and Lemma B.2 we have $\varphi_j(x;q',n) \leq x \leq \varphi_j(x;q'',n)$ with at least one inequality holding strictly. Now consider $j < nq'+1$ or $j > nq''+1$. In this case it follows from Lemma B.3 that $\tilde{g}_j(\cdot|x;q',n) \succ_{LR} \tilde{g}_j(\cdot|x;q'',n)$ so that $\varphi_j(x;q',n) < \varphi_j(x;q'',n)$ holds as a standard implication of likelihood ratio dominance.

To show part (3), suppose that G is strictly log-concave so that $\frac{g(\cdot)}{G(\cdot)}$ is strictly increasing. By Lemmas B.1 and B.2, for $j < nq+1$ we have

$$\varphi_j'(x;q,n) = \int_0^1 \underline{t}_x(x,y)\tau_{nq-1}\left(y;\frac{j-1}{nq}\right)dy < \int_0^1 \tau_{nq-1}\left(y;\frac{j-1}{nq}\right)dy = 1$$

The second step follows from part (3) of Lemma B.2. The proof for part (4) is analogous. □

B.1 Relevant properties of $\phi_n(\cdot)$ for finite n when $\rho > 0$

In this subsection we establish Proposition B.2 below, which summarizes relevant properties of $\phi_n(\cdot)$ when $\rho > 0$. The uniform Liptschitz continuity properties in statement (1) of this proposition shall play important roles in the proofs of Lemmas 2 and 5 below. Statements (2) to (4) of this proposition are consequences of the inference based on the pivotal voter's choice explained in Section 4. The third and fourth statements say that the indifference curve $\phi_n(\cdot)$ systematically shifts downwards – resulting in a preference shift towards the reform – as q increases or as the weighting function $w(\cdot)$ decreases in the sense of first order stochastic dominance. These properties play crucial roles in establishing the comparative static results in Section 6.2.

Proposition B.2. *Suppose $\rho > 0$. Then $\phi_n(\cdot)$ satisfies the following properties:*

1. *$\phi_n(\cdot)$, $\phi_n'(\cdot)$ and $\phi_n''(\cdot)$ are L-Liptschitz continuous on $[\underline{v},\bar{v}]$ for all $n \geq 0$ and some $L > 0$.*
2. *$\phi_n(x)$ is strictly increasing in x.*
3. *For any $x \in (\underline{v},\bar{v})$, $\phi_n(x)$ is strictly decreasing in q.*
4. *For any $x \in (\underline{v},\bar{v})$, $\phi_n(x)$ is weakly decreasing as $w(\cdot)$ shifts from $w^I(\cdot)$ to $w^{II}(\cdot)$, where $w^I(\cdot), w^{II}(\cdot) \in \Delta([-1,1])$ and $w^I(\cdot)$ is first order stochastically dominated by $w^{II}(\cdot)$.*

Proof of Proposition B.2. We first show part (1). Note that $\phi_n(\cdot)$ is three times continuous differentiable. By the Mean Value Theorem, $\forall x,y \in [\underline{v},\bar{v}]$ we have $|\phi_n(x)-\phi_n(y)| = |x-y|\cdot|\phi_n'(\xi)|$ for some ξ between x and y. Notice that

$$\left|\phi_n'(\xi)\right| = \rho \left|\sum_{j=1}^{n+1} w_j \varphi_j'(\xi;q,n)\right| \le \max_{j\in\{1,\cdots,n+1\}} \left|\varphi_j'(\xi;q,n)\right|$$

By (B.3), each $\varphi_j'(\xi;q,n)$ is the expectation of either $\underline{t}_x(\xi,\cdot)$ or $\bar{t}_x(\xi,\cdot)$ under some distribution. Because both $\underline{t}_x(\cdot)$ and $\bar{t}_x(\cdot)$ are uniformly bounded, there exists $L>0$ such that $L \ge \max\{\underline{t}_x(x,y), \bar{t}_x(x,y)\}$ for all $(x,y) \in [\underline{v},\bar{v}]\times[0,1]$. These together imply

$$|\phi_n(x) - \phi_n(y)| = |x-y|\cdot\left|\phi_n'(\xi)\right| < L\cdot|x-y|$$

for all $x,y \in [\underline{v},\bar{v}]$ and $n \ge 0$. The proofs for uniform Lipschitz continuities for $\phi_n'(\cdot)$ and $\phi_n''(\cdot)$ follow from analogous arguments by exploiting uniform boundedness of $\underline{t}_{xx}(\cdot)$, $\bar{t}_{xx}(\cdot)$, $\underline{t}_{xxx}(\cdot)$ and $\bar{t}_{xxx}(\cdot)$.

Next we show parts (2) and (3). Recall that

$$\phi_n(x) = \rho \sum_{j=1}^{n+1} w_j \varphi_j(x;q,n) + (1-\rho)\chi$$

By Proposition B.1, for all $j = 1,\cdots,n+1$ it holds that $\varphi_j(x;q,n)$ is strictly increasing in x and decreasing in q. Therefore, $\phi_n(x)$ must inherit these properties whenever $\rho > 0$. This proves parts (2) and (3). To show part (4), note that

$$\begin{aligned}\sum_{j=1}^{n+1} w_j \varphi_j(x;q,n) &= \sum_{j=2}^{n+1}\left[\sum_{l=j}^{n+1} w_l\right]\left(\varphi_j(x;q,n) - \varphi_{j-1}(x;q,n)\right) + \varphi_1(x;q,n) \\ &= \sum_{j=2}^{n+1}\left[1 - w\left(\frac{j-1}{n+1}\right)\right]\left(\varphi_j(x;q,n) - \varphi_{j-1}(x;q,n)\right) + \varphi_1(x;q,n)\end{aligned}$$

Consider two weighting functions $w^{\mathrm{I}}(\cdot)$ and $w^{\mathrm{II}}(\cdot)$. Let $\phi_n^{\mathrm{I}}(\cdot)$ and $\phi_n^{\mathrm{II}}(\cdot)$ denote function $\phi_n(\cdot)$ when $w(\cdot)$ equals $w^{\mathrm{I}}(\cdot)$ and $w^{\mathrm{II}}(\cdot)$, respectively. Using the above equation we obtain

$$\phi_n^{\mathrm{I}}(x) - \phi_n^{\mathrm{II}}(x) = \rho \sum_{j=2}^{n+1}\left[w^{\mathrm{II}}\left(\frac{j-1}{n+1}\right) - w^{\mathrm{I}}\left(\frac{j-1}{n+1}\right)\right]\left(\varphi_j(x;q,n) - \varphi_{j-1}(x;q,n)\right)$$

By Proposition B.1, $\varphi_j(x;q,n) - \varphi_{j-1}(x;q,n) > 0$ holds for all $j > 1$ and $x \in (\underline{v},\bar{v})$. Suppose $w^{\mathrm{II}}(\cdot)$ first order stochastically dominates $w^{\mathrm{I}}(\cdot)$, then $w^{\mathrm{II}}(y) - w^{\mathrm{I}}(y) \le 0$ holds for all $y \in (0,1)$. This implies $\phi_n^{\mathrm{I}}(x) - \phi_n^{\mathrm{II}}(x) \le 0$ for all $n \ge 0$ and $x \in (\underline{v},\bar{v})$. □

B.2 Asymptotic properties of $\phi_n(\cdot)$ and the proofs of Lemmas 2 and 3

In this subsection we derive asymptotic properties of $\phi_n(\cdot)$ as $n \to \infty$ and prove Lemmas 2 and 3 in Section 4. Moreover, we also establish some additional properties in Lemmas B.4 and B.5 below; these are relevant for proofs in Appendices D and E.

Given a sender's preference parameters ρ, $w(\cdot)$ and χ, we define

$$\phi(x) := \rho \left[\int_0^q \underline{t}\left(x, \frac{y}{q}\right) dw(y) + \int_q^1 \bar{t}\left(x, \frac{y-q}{1-q}\right) dw(y) \right] + (1-\rho)\chi \tag{B.8}$$

for $x \in [\underline{v}, \bar{v}]$, where $\underline{t}(\cdot)$ and $\bar{t}(\cdot)$ are given (B.4) and (B.5), respectively. The first order derivative of $\phi(x)$ is given by

$$\phi'(x) = \rho \left[\int_0^q \underline{t}_x\left(x, \frac{y}{q}\right) dw(y) + \int_q^1 \bar{t}_x\left(x, \frac{y-1}{1-q}\right) dw(y) \right] \tag{B.9}$$

Moreover, using (B.4), (B.5) and the fact that $v_q^* = G^{-1}(q)$, we obtain

$$\underline{t}\left(v_q^*, \frac{y}{q}\right) = G^{-1}\left(\frac{y}{q} G(v_q^*)\right) = G^{-1}(y)$$

and

$$\bar{t}\left(v_q^*, \frac{y-q}{1-q}\right) = G^{-1}\left(\frac{y-q}{1-q} + (1 - \frac{y-q}{1-q}) G(v_q^*)\right) = G^{-1}(y)$$

These together imply

$$\phi^* := \phi(v_q^*) = \rho \int_0^1 G^{-1}(y) dw(y) + (1-\rho)\chi \tag{B.10}$$

Lemma 2 in Section 4 is then equivalent to that $\phi_n(\cdot)$ and $\phi_n'(\cdot)$ uniformly converge to (B.8) and (B.9), respectively, and $\varphi_n(v)$ converges almost surely to (B.10).

B.2.1 Proof of Lemma 2

Consider any $z \in (0,1)$. By (B.3) in Lemma B.1 we have

$$\varphi_{\lfloor (n+1)z \rfloor}(x; q, n) = \begin{cases} \int_0^1 \underline{t}(x,y) \tau_{nq-1}\left(y; \frac{\lfloor (n+1)z \rfloor - 1}{nq}\right) dy, & \text{if } z < \frac{nq+1}{n+1} \\ \int_0^1 \bar{t}(x,y) \tau_{n(1-q)-1}\left(y; \frac{\lfloor (n+1)z \rfloor - nq - 2}{n(1-q)}\right) dy, & \text{if } z > \frac{nq+1}{n+1} \end{cases}$$

By Lemma A.1c, as $n \to \infty$, $\tau_{nq-1}\left(y; \frac{\lfloor (n+1)z \rfloor - 1}{nq}\right)$ concentrates all its probability mass on $\frac{z}{q}$ and $\tau_{n(1-q)-1}\left(y; \frac{\lfloor (n+1)z \rfloor - nq - 2}{n(1-q)}\right)$ concentrates all its mass on $\frac{z-q}{1-q}$ for all $z \neq q$. Therefore,

$$\lim_{n\to\infty} \varphi_{\lfloor (n+1)z \rfloor}(x; q, n) = \begin{cases} \underline{t}\left(x, \frac{z}{q}\right), & \text{if } z < q \\ \bar{t}\left(x, \frac{z-q}{1-q}\right), & \text{if } z > q \end{cases} \tag{B.11}$$

Using the definition of $\phi_n(x)$ and the fact that $w_j = w\left(\frac{j}{n+1}\right) - w\left(\frac{j-1}{n+1}\right)$, we get

$$\phi_n(x) = \rho \sum_{j=1}^{n+1} \left[w(z_j) - w(z_{j-1})\right] \varphi_{(n+1)z_j}(x; q, n) + (1-\rho)\chi \tag{B.12}$$

where $z_j := \frac{j}{n+1}$. Taking $n \to \infty$ and using the definition of a Riemann integral, we have

$$\begin{aligned} \lim_{n\to\infty} \sum_{j=1}^{n+1} \left[w(z_j) - w(z_{j-1})\right] \varphi_{(n+1)z_j}(x; q, n) &= \int_0^1 \lim_{n\to\infty} \varphi_{\lfloor (n+1)z \rfloor}(x; q, n) dw(z) \\ &= \int_0^q \underline{t}\left(x, \frac{y}{q}\right) dw(y) + \int_q^1 \bar{t}\left(x, \frac{y-q}{1-q}\right) dw(y) \end{aligned}$$

where the last step follows from (B.11). Combining this with (B.12) yields

$$\lim_{n\to\infty} \phi_n(x) = \rho \left[\int_0^q \underline{t}\left(x, \frac{y}{q}\right) dw(y) + \int_q^1 \bar{t}\left(x, \frac{y-q}{1-q}\right) dw(y)\right] + (1-\rho)\chi \tag{B.13}$$

Therefore, $\phi_n(x)$ converges point-wise to (B.8) on $[\underline{v}, \bar{v}]$. By statement (1) of Proposition B.2, $\phi_n(x)$ is uniformly L-Liptschitz continuous on $[\underline{v}, \bar{v}]$ for some sufficient large L. Following the same argument in the proof of Proposition B.2, it can be show that $\phi(x)$ given by (B.8) is also L-Liptschitz continuous for sufficiently large L. These together imply that the convergence of $\phi_n(x)$ to (B.8) is in fact uniform.[2] The fact that $\phi_n'(x)$ converges uniformly to (B.9) can be proved analogously.

[2] This stems from the following general observation: *For any $L > 0$, let $\{f_n(\cdot)\}_{n\geq 0}$ be a sequence of L-Lipschitz continuous functions on a closed interval $[a,b]$ that converges point-wise to a L-Lipschitz continuous function $f(\cdot)$. Then $f_n(\cdot)$ converges uniformly to $f(\cdot)$ on $[a,b]$.* The proof is as follows. Given any $\varepsilon > 0$, consider a pair of $\delta, \eta > 0$ such that $\varepsilon > 2L\delta + \eta$. Partition interval $[a,b]$ into $K+1$ intervals with cutoffs $a = x_0 < x_1 < \cdots < x_{K+1} = b$ such that $|x_i - x_{i-1}| < \delta$ for all $i \in \{1, \cdots, K+1\}$. For this finite set $\{x_i\}_{i=1}^{K+1}$ point-wise convergence implies that there exists a threshold N such that for all $n > N$ we have $|f_n(x_i) - f(x_i)| < \eta$ for all $i \in \{1, \cdots, K\}$. Now consider any $x \in [a,b]$ and let i be such that $x \in [x_{i-1}, x_i]$. We then obtain

$$\begin{aligned} |f_n(x) - f(x)| &= |f_n(x) - f_n(x_i) + f_n(x_i) - f(x_i) + f(x_i) - f(x)| \\ &\leq |f_n(x) - f_n(x_i)| + |f_n(x_i) - f(x_i)| + |f(x_i) - f(x)| \\ &< 2L\delta + \eta < \varepsilon \end{aligned}$$

for all $n > N$. This proves uniform convergence on $[a,b]$.

Finally, we prove that $\varphi_n(v) := \rho \sum_{j=1}^{n+1} w_j v^{(j)} + (1-\rho)\chi$ converges in probability to ϕ^* given by (B.10). It suffices to show that $\sum_{j=1}^{n+1} w_j v^{(j)}$ converges in probability to $\int_0^1 G^{-1}(y)dw(y)$. To prove this, we use a strong law of large numbers for linear combinations of order statistics established by Van Zwet (1980). Let $U_1, U_2, \cdots, U_{n+1}$ be $n+1$ random variables drawn from a uniform distribution on $(0,1)$, and $U_{1:n+1} \leq U_{2:n+1} \leq \cdots \leq U_{n+1:n+1}$ denote the ordered $U_1, U_2, \cdots, U_{n+1}$. We can therefore rewrite $v^{(j)}$ as $G^{-1}(U_{j:n+1})$ for each $j = 1, \cdots, n+1$. For $t \in (0,1)$, define $\gamma_n(t) := G^{-1}(U_{\lceil (n+1)t \rceil : n+1})$ and $\xi_n(t) := (n+1) \cdot \left[w\left(\frac{\lceil (n+1)t \rceil}{n+1}\right) - w\left(\frac{\lceil (n+1)t \rceil - 1}{n+1}\right)\right]$. We can then rewrite $\sum_{j=1}^{n+1} w_j v^{(j)}$ in an integral form as

$$\sum_{j=1}^{n+1} w_j v^{(j)} = \sum_{j=1}^{n+1} \left[w\left(\frac{j}{n+1}\right) - w\left(\frac{j-1}{n+1}\right)\right] G^{-1}(U_{j:n+1}) = \int_0^1 \gamma_n(t)\xi_n(t)dt \tag{B.14}$$

Our assumptions for G and $w(\cdot)$ ensure that $G^{-1}(\cdot), w(\cdot) \in L_1$, $\sup_n \|\xi_n\|_\infty < \infty$, and $\lim_{n\to\infty} \int_0^t \xi_n(x)dx = \int_0^t w'(x)dx = w(t)$ for all $t \in (0,1)$.[3] It follows from Theorem 2.1 and Corollary 2.1 of Van Zwet (1980) that the integral in (B.14) converges almost surely to $\int_0^1 G^{-1}(y)dw(y)$ as $n \to \infty$.

B.2.2 Proof of Lemma 3

In case n is finite, it follows from (B.1) that

$$\phi_n'(x) = \rho \sum_{j=1}^{n+1} w_j \varphi_j'(x; q, n) \tag{B.15}$$

Because $\varphi_j'(x;q,n)$ is uniformly bounded for all j, $\phi_n'(x)$ converges uniformly to zero as $\rho \to 0$. Therefore, there exists $\overline{\rho} > 0$ such that $\phi_n'(x) \leq 1$ for all $x \in [\underline{v}, \overline{v}]$ and $\rho \leq \overline{\rho}$. Moreover, by Proposition B.1, if both G and $1-G$ are strictly log-concave, then $\varphi_j'(x;q,n) < 1$ for all $j \neq nq+1$ and $\varphi_j'(x;q,n) = 1$ for $j = nq+1$. These together imply ${\phi_n}'(x) \leq \rho \leq 1$ for all $x \in [\underline{v}, \overline{v}]$ by (B.15).

In what follows we show for all $x \in [\underline{v}, \overline{v}]$ that $\phi'(x) < 1$ holds under either condition (i) and (ii) of Lemma 3. Because both $\underline{t}_x\left(x, \frac{y}{q}\right)$ and $\overline{t}_x\left(x, \frac{y-1}{1-q}\right)$ are uniformly bounded, it follows from (B.9) that $\phi'(x)$ must uniformly converge to zero as $\rho \to 0$. This implies $\phi'(x) < 1$ for all $x \in [\underline{v}, \overline{v}]$ if ρ is sufficiently close to zero. If instead both G and $1-G$ are strictly log-concave, it follows from Lemma B.2 that $\underline{t}_x\left(x, \frac{y}{q}\right) < 1$ for $y < q$ and $\overline{t}_x\left(x, \frac{y-q}{1-q}\right) < 1$ for $y > q$. Therefore, by (B.9), $\phi'(x) < 1$ holds uniformly on $[\underline{v}, \overline{v}]$ if either $\rho < 1$ or $w(\cdot)$ does not put all weights on $y = q$ (i.e., $w(\cdot)$ is not a step function with threshold q). The latter condition for $w(\cdot)$ is always satisfied due to our assumption that $w(\cdot)$ is absolutely continuous.

[3] Here L_1 refers to the space of Lebesgue measurable functions $f : (0,1) \mapsto \mathbb{R}$ with finite $\|\cdot\|_1$ norm. $\|\cdot\|_\infty$ denotes the essential supremum norm. Moreover, the absolute continuity of $w(\cdot)$ ensures that its derivative $w'(\cdot)$ exists almost everywhere and $\int_0^t w'(x)dx = w(t)$ for all $t \in (0,1)$.

B.2.3 Additional useful properties of $\phi(\cdot)$ and ϕ^*

Here we establish several additional properties for $\phi(\cdot)$ and ϕ^* summarized in Lemmas B.4 and B.5. The results are relevant for the proofs in Appendix E.

Lemma B.4. *For any $x \in [\underline{v}, \bar{v}]$, the following properties hold:*

1. If $\rho > 0$, then $\phi(x)$ is strictly decreasing in q.

2. If $\rho > 0$, then $\phi(x)$ strictly decreases as $w(\cdot)$ shifts from $w^I(\cdot)$ to $w^{II}(\cdot)$, where $w^I(\cdot), w^{II}(\cdot) \in \Delta([-1,1])$ and $w^I(\cdot) \succeq_{FOSD} w^{II}(\cdot)$.[4]

Proof of Lemma B.4. For all $y \in [0,1]$ let

$$t(x,y;q) := \begin{cases} \underline{t}\left(x, \frac{y}{q}\right), & \text{if } y \leq q \\ \bar{t}\left(x, \frac{y-q}{1-q}\right), & \text{if } y > q \end{cases} \tag{B.16}$$

where $\underline{t}(\cdot)$ and $\bar{t}(\cdot)$ are given by (B.4) and (B.5), respectively. By (B.8) we can rewrite $\phi(\cdot)$ as

$$\phi(x) = \rho \int_0^1 t(x,y;q)dw(y) + (1-\rho)\chi = \rho \mathbb{E}_w[t(x,\cdot;q)] + (1-\rho)\chi \tag{B.17}$$

Notice that $\phi(x)$ depends on x, q and $w(\cdot)$ only through the integral $\int_0^1 t(x,y;q)dw(y)$. Because both $\underline{t}(x,y)$ and $\bar{t}(x,y)$ are strictly increasing in y (cf. Lemma B.2), it follows from (B.16) that $t(x,y;q)$ is strictly decreasing in q. $\int_0^1 t(x,y;q)dw(y)$ must inherit the same property and therefore part (1) holds. Next, consider two weighting functions $w^{\text{I}}(\cdot)$ and $w^{\text{II}}(\cdot)$ such that $w^{\text{I}}(\cdot) \succeq_{FOSD} w^{\text{II}}(\cdot)$. Because $t(x,y;q)$ is strictly increasing in y, $\int_0^1 t(x,y;q)dw^{\text{I}}(y) > \int_0^1 t(x,y;q)dw^{\text{II}}(y)$ must hold. Therefore, $\phi(x)$ must be strictly higher under $w^{\text{I}}(\cdot)$ than under $w^{\text{II}}(\cdot)$. This proves part (2). □

Building on Lemma B.4, we establish our next Lemma B.5, which characterizes how ϕ^* varies with model primitives $w(\cdot)$ and q.

Lemma B.5. *Suppose $\rho > 0$. Then ϕ^* is invariant in q and it strictly decreases as $w(\cdot)$ shifts from $w^I(\cdot)$ to $w^{II}(\cdot)$, where $w^I(\cdot) \succeq_{FOSD} w^{II}(\cdot)$.*

Proof of Lemma B.5. Recall that

$$\phi^* = \phi(v_q^*) = \rho \int_0^1 G^{-1}(y)dw(y) + (1-\rho)\chi$$

It is clear from its expression that ϕ^* is invariant in q. Consider any $w^{\text{I}}(\cdot)$ and $w^{\text{II}}(\cdot)$ with $w^{\text{I}}(\cdot) \succeq_{FOSD} w^{\text{II}}(\cdot)$. Then $\int_0^1 G^{-1}(y)dw^{\text{I}}(y) > \int_0^1 G^{-1}(y)dw^{\text{II}}(y)$ must hold because $G^{-1}(y)$ is strictly increasing. Since $\rho > 0$, it follows that ϕ^* strictly decreases as $w(\cdot)$ shifts from $w^{\text{I}}(\cdot)$ to $w^{\text{II}}(\cdot)$. □

[4] We use notation $\succeq_{FOSD}$ to denote the partial order implied by first order stochastic dominance.

C Omitted proofs for Section 5

C.1 Proof of Lemma 4

Our proof for Lemma 4 proceeds in two steps. In Step 1 we establish a general property (cf. Observation C.1) that a sender's utility function satisfying the increasing-slope property at any interior point z implies $H \succeq_{MPS} H_{\mathscr{P}(z)}$ for any H that solves his monopolistic persuasion problem. In Step 2 we subsequently show that $W_n(\cdot)$ indeed satisfies the increasing slope property at switching point z_n whenever it is interior in $(-1,1)$. These together establish our Lemma 4.

Step 1. Let $U(\cdot)$ be a generic utility function (of voters' posterior expected state θ) defined on $[-1,1]$. Then, for any prior $F \in \Delta([-1,1])$ (which need not be continuous and fully supported), let

$$\mathscr{U}_\pi(U,F) := \mathbb{E}_{H_\pi}[U(\cdot)] = \int_{-1}^{1} U(\theta) dH_\pi(\theta)$$

denote the sender's expected payoff under any feasible information policy $\pi \in \Pi$.[5] Let $\underline{\pi}$ denote the *null* information policy that reveals no information. Then $H_{\underline{\pi}}$ is a degenerate distribution with all mass on prior mean $\mu_F := \mathbb{E}_F[k]$ and therefore

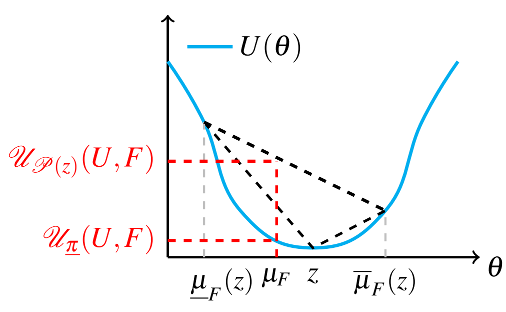
$$\mathscr{U}_{\underline{\pi}}(U,F) = U(\mu_F) \tag{C.1}$$

On the other hand, for a cutoff censorship policy $\mathscr{P}(z)$ with $z \in (-1,1)$, we have

$$\mathscr{U}_{\mathscr{P}(z)}(U,F) := F^-(z)U(\underline{\mu}_F(z)) + \big(F(z) - F^-(z)\big)\, U(z) + (1 - F(z))\, U(\overline{\mu}_F(z)) \tag{C.2}$$

where $F^-(z) := \lim_{x \uparrow z} F(z)$, $\underline{\mu}_F(z) := \mathbb{E}_F[k|k < z]$ and $\overline{\mu}_F(z) := \mathbb{E}_F[k|k > z]$. Figure C.1 illustrates $\mathscr{U}_{\mathscr{P}(z)}(U,F)$ and $\mathscr{U}_{\underline{\pi}}(U,F)$ for a function $U(\cdot)$ that satisfies the increasing slope property at some $z \in (-1,1)$ and a prior F with no mass point at z.

Figure C.1: $\mathscr{U}_{\mathscr{P}(z)}(U,F)$ and $\mathscr{U}_{\underline{\pi}}(U,F)$



[5] Recall that H_π denotes the distribution of posterior expectations induced by π under prior F.

Claim C.1. *Suppose $U(\cdot)$ satisfies the increasing slope property at some point $z \in (-1,1)$, then $\mathscr{U}_{\mathscr{P}(z)}(U,F) > \mathscr{U}_{\underline{\pi}}(U,F)$ for any $F \in \Delta([-1,1])$ that satisfies $0 < F^-(z) \leq F(z) < 1$.*

Proof of Claim C.1. By (C.1) and (C.2), we obtain

$$\begin{aligned}\mathscr{U}_{\mathscr{P}(z)}(U,F) - \mathscr{U}_{\underline{\pi}}(U,F) =& F^-(z)\left(\underline{\mu}_F(z) - \mu_F\right)\frac{U(\underline{\mu}_F(z)) - U(\mu_F)}{\underline{\mu}_F(z) - \mu_F} \\ &+ \left(F(z) - F^-(z)\right)(z - \mu_F)\frac{U(z) - U(\mu_F)}{z - \mu_F} \\ &+ (1 - F(z))\left(\overline{\mu}_F(z) - \mu_F\right)\frac{U(\overline{\mu}_F(z)) - U(\mu_F)}{\overline{\mu}_F(z) - \mu_F}\end{aligned}$$

On the other hand, by the law of iterated expectations, we have

$$\begin{aligned}& F^-(z)\underline{\mu}_F(z) + \left(F(z) - F^-(z)\right)z + (1 - F(z))\overline{\mu}_F(z) = \mu_F \\ \Longrightarrow\ & \left(F(z) - F^-(z)\right)(z - \mu_F) = -F^-(z)\left(\underline{\mu}_F(z) - \mu_F\right) - (1 - F(z))\left(\overline{\mu}_F(z) - \mu_F\right)\end{aligned}$$

These together imply

$$\begin{aligned}\mathscr{U}_{\mathscr{P}(z)}(U,F) - \mathscr{U}_{\underline{\pi}}(U,F) = & F^-(z)\left(\mu_F - \underline{\mu}_F(z)\right)\left(\frac{U(z) - U(\mu_F)}{z - \mu_F} - \frac{U(\mu_F) - U(\underline{\mu}_F(z))}{\mu_F - \underline{\mu}_F(z)}\right) \\ &+ (1 - F(z))\left(\overline{\mu}_F(z) - \mu_F\right)\left(\frac{U(\overline{\mu}_F(z)) - U(\mu_F)}{\overline{\mu}_F(z) - \mu_F} - \frac{U(z) - U(\mu_F)}{z - \mu_F}\right)\end{aligned}$$

Since $\underline{\mu}_F(z) < x < \overline{\mu}_F(z)$ for $x \in \{z, \mu_F\}$, the increasing slope property at z implies $\frac{U(z)-U(\mu_F)}{z-\mu_F} - \frac{U(\mu_F)-U(\underline{\mu}_F(z))}{\mu_F-\underline{\mu}_F(z)} \geq 0$ and $\frac{U(\overline{\mu}_F(z))-U(\mu_F)}{\overline{\mu}_F(z)-\mu_F} - \frac{U(z)-U(\mu_F)}{z-\mu_F} \geq 0$, with at least one of these holding with strict inequality.[6] This implies $\mathscr{U}_{\mathscr{P}(z)}(U,F) - \mathscr{U}_{\underline{\pi}}(U,F) \geq 0$ for all F. Finally, notice that if $0 < F^-(z) \leq F(z) < 1$ holds, then both $F^-(z)\left(\mu_F - \underline{\mu}_F(z)\right)$ and $(1 - F(z))\left(\overline{\mu}_F(z) - \mu_F\right)$ are strictly positive so that $\mathscr{U}_{\mathscr{P}(z)}(U,F) - \mathscr{U}_{\underline{\pi}}(U,F) > 0$ must hold. □

We are now ready to establish the following general observation.

Observation C.1. *Suppose $U(\cdot)$ satisfies the increasing slope property at point $z \in (-1,1)$. Then $H \succeq_{MPS} H_{\mathscr{P}(z)}$ for any H (if it exists) that solves*

$$\max_{H \in \Delta([-1,1])} \int_{-1}^{1} U(\theta)dH(\theta), \quad s.t.\ F \succeq_{MPS} H \tag{C.3}$$

[6] In case $z = \mu_F$, we can let $\frac{U(z)-U(\mu_F)}{z-\mu_F}$ be any number between $U'_-(z)$ and $U'_+(z)$, which are the left and right derivatives of $U(\cdot)$ at point z, respectively. Notice that the increasing slope property at point z implies the existence of both $U'_-(z)$ and $U'_+(z)$ (through the monotone convergence theorem) and that $U'_-(z) \leq U'_+(z)$.

Proof of Observation C.1. Suppose $U(\cdot)$ satisfies the increasing slope property at point z and let H be any solution to (C.3) (if it exists). We show by contradiction that $H \succeq_{MPS} H_{\mathscr{P}(z)}$ must hold. Suppose there exists any $H \not\succeq_{MPS} H_{\mathscr{P}(z)}$ that solves (C.3). Let $\pi = (S, \sigma)$ be an information policy that induces H. For each $s \in S$, let γ_s denote the posterior distribution induced by s and h_s denote the mean of γ_s. Finally, let $\delta \in \Delta(S)$ denote the ex-ante distribution of messages $s \in S$ induced by π. With these we obtain

$$\int_{-1}^{1} U(\theta) dH(\theta) = \int_{s \in S} U(h_s) d\delta(s) = \int_{s \in S} \mathscr{U}_{\underline{\pi}}(U, \gamma_s) d\delta(s)$$

Since $H \not\succeq_{MPS} H_{\mathscr{P}(z)}$, there exists $s \in S$ such that $0 < \gamma_s^-(z) \leq \gamma_s(z) < 1$ holds. Denote by $\widetilde{S} \subseteq S$ the set of all such s. Then $\widetilde{S}$ must have positive probability measure under δ. Consider the joint information policy $\widetilde{\pi}$ induced by π and the cutoff policy $\mathscr{P}(z)$, and let $\widetilde{H} = H_{\widetilde{\pi}}$. For all events $s \in \widetilde{S}$, it follows from Claim C.1 that $\mathscr{U}_{\mathscr{P}(z)}(U, \gamma_s) > \mathscr{U}_{\underline{\pi}}(U, \gamma_s)$. For $s \notin \widetilde{S}$, $\mathscr{U}_{\mathscr{P}(z)}(U, \gamma_s) = \mathscr{U}_{\underline{\pi}}(U, \gamma_s)$ holds trivially. Therefore, we have

$$\begin{aligned} \int_{-1}^{1} U(\theta) d\widetilde{H}(\theta) &= \int_{s \in \widetilde{S}} \mathscr{U}_{\mathscr{P}(z)}(U, \gamma_s) d\delta(s) + \int_{s \in S/\widetilde{S}} \mathscr{U}_{\mathscr{P}(z)}(U, \gamma_s) d\delta(s) \\ &= \int_{s \in \widetilde{S}} \mathscr{U}_{\mathscr{P}(z)}(U, \gamma_s) d\delta(s) + \int_{s \in S/\widetilde{S}} \mathscr{U}_{\underline{\pi}}(U, \gamma_s) d\delta(s) \qquad \text{(C.4)} \\ &> \int_{s \in \widetilde{S}} \mathscr{U}_{\underline{\pi}}(U, \gamma_s) d\delta(s) + \int_{s \in S/\widetilde{S}} \mathscr{U}_{\underline{\pi}}(U, \gamma_s) d\delta(s) = \int_{-1}^{1} U(\theta) dH(\theta) \end{aligned}$$

This contradicts that H is a solution to (C.3) and thus completes the proof. □

Step 2. Now we establish that $W_n(\cdot)$ satisfies the increasing slope property at z_n when $z_n \in (-1, 1)$ holds. Recall from (9) in Section 5 that

$$W_n(\theta) := \int_{\underline{v}}^{\theta} (\theta - \phi_n(x))\, \hat{g}_n(x;q) dx = \theta \hat{G}_n(\theta;q) - \int_{\underline{v}}^{\theta} \phi_n(x) \hat{g}_n(x;q) dx \qquad \text{(C.5)}$$

By (C.5), for any $\theta \neq z_n$ we have

$$\begin{aligned} W_n(\theta) - W_n(z_n) &= \int_{\underline{v}}^{\theta} (\theta - \phi_n(x))\, \hat{g}_n(x;q) dx - \int_{\underline{v}}^{z_n} (z_n - \phi_n(x))\, \hat{g}_n(x;q) dx \\ &= \int_{z_n}^{\theta} (\theta - \phi_n(x))\, \hat{g}_n(x;q) dx + (\theta - z_n) \int_{\underline{v}}^{z_n} \hat{g}_n(x;q) dx \end{aligned}$$

Therefore,

$$\lambda_n(\theta; z_n) := \frac{W_n(\theta) - W_n(z_n)}{\theta - z_n} = \int_{z_n}^{\theta} \frac{\theta - \phi_n(x)}{\theta - z_n} \hat{g}_n(x;q) dx + \int_{\underline{v}}^{z_n} \hat{g}_n(x;q) dx$$

Taking the derivative with respect to θ yields

$$\lambda_n{}'(\theta;z_n) = \frac{\theta-\phi_n(\theta)}{\theta-z_n}\hat{g}_n(\theta;q) + \int_{z_n}^{\theta} \frac{\phi_n(x)-z_n}{(\theta-z_n)^2}\hat{g}_n(x;q)dx \tag{C.6}$$

Recall from the premise of this lemma that $\phi_n(\cdot)$ crosses zero only once and from above at z_n. For any $\theta > z_n$, $x > \phi_n(x) \geq z_n$ holds for all $x \in (z_n, \theta]$. Therefore, the first term on the right-hand side of (C.6) must be strictly positive and the second term is non-negative. This implies $\lambda_n{}'(\theta;z_n) > 0$ for $\theta > z_n$. For any $\theta < z_n$, $x < \phi_n(x) \leq z_n$ holds for all $x \in [\theta, z_n)$. So the first term on the right-hand side of (C.6) is strictly positive, and the second term equals

$$\int_{z_n}^{\theta} \frac{\phi_n(x)-z_n}{(\theta-z_n)^2}\hat{g}_n(x;q)dx = \int_{\theta}^{z_n} \frac{z_n-\phi_n(x)}{(\theta-z_n)^2}\hat{g}_n(x;q)dx$$

and is non-negative. This implies $\lambda_n{}'(\theta;z_n) > 0$ for $\theta < z_n$ as well. Taken together, $\lambda_n{}'(\theta;z_n) > 0$ holds for all $\theta \neq z_n$. Finally, since $W_n(\theta)$ is differentiable, we have that $\lambda_n(\theta;z_n)$ is continuous at $\theta = z_n$ and $\lim_{\theta\to z_n} \lambda_n(\theta;z_n) = W_n{}'(z_n)$. These together establish that $\lambda_n(\theta;z_n)$ is strictly increasing in θ, which implies the increasing slope property at point z_n. Together with Observation C.1, this establishes our Lemma 4.

C.2 Proof of Lemma 5

By (9), the second order derivative of $W_n(\theta)$ is given by

$$\begin{aligned} W_n''(\theta) &= \hat{g}_n(\theta;q)\left(2-\phi_n'(\theta)\right) + \hat{g}_n'(\theta;q)\left(\theta-\phi_n'(\theta)\right) \\ &= \hat{g}_n(\theta;q)\left\{2-\phi_n{}'(\theta) + (\theta-\phi_n(\theta))\frac{\hat{g}_n'(\theta;q)}{\hat{g}_n(\theta;q)}\right\} \\ &= \hat{g}_n(\theta;q)\left\{2-\phi_n'(\theta) + (\theta-\phi_n(\theta))\left(n\frac{g(\theta)}{G(\theta)}\frac{q-G(\theta)}{1-G(\theta)} + \frac{g'(\theta)}{g(\theta)}\right)\right\} \end{aligned} \tag{C.7}$$

The last step follows from (A.9). Because G is strictly positive and twice-continuously differentiable on $[\underline{v}, \overline{v}]$, $W_n''(\theta)$ is continuous. By (C.7) and the fact that $q = G(v_q^*)$, $W_n''(\theta) > 0$ if and only if

$$(\theta-\phi_n(\theta))\left(G(\theta)-G(v_q^*)\right) < \frac{G(\theta)\left(1-G(\theta)\right)}{ng(\theta)}\left(2-\phi_n{}'(\theta) + (\theta-\phi_n(\theta))\frac{g'(\theta)}{g(\theta)}\right) \tag{C.8}$$

Because $\phi_n(\cdot)$, $\phi_n'(\cdot)$ and $\phi_n''(\cdot)$ are uniformly Lipschitz continuous (cf. Proposition B.2) and $g(\cdot)$ is positive and twice continuously differentiable on $[\underline{v}, \overline{v}]$, both the value and the first order derivative of the right-hand side of (C.8) converge to zero uniformly for all $\theta \in [\underline{v}, \overline{v}]$.[7] Let

[7] This is because both $1-\phi_n{}'(\theta)$ and $\frac{g'(\theta)}{G'(\theta)}$ are uniformly bounded on $[\underline{v}, \overline{v}]$ under our assumption for G.

$\zeta_n(\theta) := (\theta - \phi_n(\theta))\left(G(\theta) - G(v_q^*)\right)$ denote the left-hand side of (C.8) and

$$\zeta(\theta) := \lim_{n\to\infty} \zeta_n(\theta) = (\theta - \phi(\theta))\left(G(\theta) - G(v_q^*)\right). \tag{C.9}$$

Both the value and derivative of $\zeta_n(\theta)$ converge uniformly to $\zeta(\cdot)$ and $\zeta'(\cdot)$, respectively. Therefore, $\lim_{n\to\infty} W_n''(\theta) > (<) 0$ if and only if $\zeta(\theta) < (>) 0$. On the one hand, $G(\theta) - G(v_q^*)$ is increasing in θ and admits a unique root v_q^* at which its derivative equals $g(v_q^*) > 0$. On the other hand, under our definition of the single-crossing property (cf. Definition 1), $\theta - \phi(\theta)$ crosses zero at most once and from below on $[-1,1]$. Recall that $z^* = \lim_{n\to\infty} z_n$ (cf. (12)) and the single-crossing property requires $1 - \phi'(z^*) > 0$ whenever $\phi(z^*) = z^*$ and $z^* \in [-1,1]$. Let

$$\ell^* := \max\left\{\min\left\{z^*, v_q^*\right\}, -1\right\} \quad \text{and} \quad r^* := \min\left\{\max\left\{z^*, v_q^*\right\}, 1\right\} \tag{C.10}$$

It follows from (C.9) that $\zeta(\theta) < 0$ for all $\theta \in (\ell^*, r^*)$ and $\zeta(\theta) > 0$ for all $\theta \in [-1,1]/[\ell^*, r^*]$. We distinguish between three cases.

Case 1: $z^* = -1$ and $\theta - \phi(\theta) > 0$ for all $\theta \in [-1,1]$. In this case, $\ell^* = -1$ and $\zeta(\theta) > (<) 0$ for $\theta > (<) r^*$ on $[-1,1]$. It then follows that there exists some $N \geq 0$ such that for all $n \geq N$ we have (i) $\theta - \phi_n(\theta) > 0$ for all $\theta \in [-1,1]$ so that $z_n = -1$, and (ii) there is some $r_n \in [-1,1]$ with $r_n \to r^*$ such that $W_n''(\theta) > (<) 0$ for $\theta < (>) r_n$. This implies that $W_n(\cdot)$ is strictly S-shaped on $[z_n, 1]$ with inflection point r_n for all $n \geq N$.

Case 2: $z^* = 1$ and $\theta - \phi(\theta) < 0$ for all $\theta \in [-1,1]$. In this case, $r^* = 1$ and $\zeta(\theta) < (>) 0$ for $\theta > (<) \ell^*$ on $[-1,1]$. There then exists $N \geq 0$ such that for all $n \geq N$ we have (i) $\theta - \phi_n(\theta) < 0$ for all $\theta \in [-1,1]$ so that $z_n = 1$, and (ii) there is some $\ell_n \in [-1,1]$ with $\ell_n \to \ell^*$ such that $W_n''(\theta) > (<) 0$ for $\theta > (<) \ell_n$. This implies that $W_n(\cdot)$ is strictly inverse S-shaped on $[-1, z_n]$ with inflection point ℓ_n for all $n \geq N$.

Case 3: $z^* \in [-1,1]$ and $z^* = \phi(z^*)$. Recall that the single-crossing property requires $1 - \phi'(z^*) > 0$ whenever $\phi(z^*) = z^* \in [-1,1]$. It follows that there exists some $N \geq 0$ and $\varepsilon > 0$ such that for all $n \geq N$ we have that

(i) there exists a unique $\tilde{z}_n \in \delta(z^*; \varepsilon) := (z^* - \varepsilon, z^* + \varepsilon)$ such that both $\tilde{z}_n = \phi_n(\tilde{z}_n)$ and $1 - \phi_n'(\tilde{z}_n) > 0$ hold, and

(ii) there are $\tilde{\ell}_n, \tilde{r}_n \in \mathbb{I} := [-1,1] \cup \delta(z^*; \varepsilon)$ such that $W_n''(\theta) > 0$ if $\theta \in \left(\tilde{\ell}_n, \tilde{r}_n\right)$ and $W_n''(\theta) < 0$ if $\theta \in \mathbb{I}/\left[\tilde{\ell}_n, \tilde{r}_n\right]$, where $\tilde{\ell}_n$ and $\tilde{r}_n$ satisfy $\tilde{\ell}_n \leq \tilde{r}_n$, $\tilde{\ell}_n \to \max\left\{\min\left\{z^*, v_q^*\right\}, \min\left\{-1, z^* - \varepsilon\right\}\right\}$, and $\tilde{r}_n \to \min\left\{\max\left\{z^*, v_q^*\right\}, \max\left\{1, z^* + \varepsilon\right\}\right\}$.

Let $\ell_n = \max\left\{\tilde{\ell}_n, -1\right\}$ and $r_n = \min\left\{\tilde{r}_n, 1\right\}$. It follows from (ii) that $W_n''(\theta) > 0$ if $\theta \in (\ell_n, r_n)$

and $W_n''(\theta) < 0$ if $\theta \in [-1,1]/[\ell_n, r_n]$. Moreover, $\ell_n \to \ell^*$ and $r_n \to r^*$. Finally, (i) implies

$$\begin{aligned} W_n''(\tilde{z}_n) &= \hat{g}_n(\tilde{z}_n;q)\left(2-\phi_n'(\tilde{z}_n)\right) + \hat{g}_n'(\tilde{z}_n;q)\left(\tilde{z}_n - \phi_n(\tilde{z}_n)\right) \\ &= \hat{g}_n(\tilde{z}_n;q)\left(2-\phi_n'(\tilde{z}_n)\right) > \hat{g}_n(\tilde{z}_n;q) > 0 \end{aligned}$$

Therefore $\tilde{z}_n \in \left(\tilde{\ell}_n, \tilde{r}_n\right)$ must hold. If $\tilde{z}_n \in [-1,1]$, then $\tilde{z}_n = z_n$ and we obtain $\ell_n \le z_n \le r_n$ with at least one inequality holding strictly. If $\tilde{z}_n < -1$ (resp. $\tilde{z}_n > 1$) then $-1 = z_n = \ell_n \le r_n$ (resp. $1 = z_n = r_n \ge \ell_n$). Taken together, for all $n \ge N$, $W_n(\cdot)$ is S-shaped on $[z_n, 1]$ with inflection point $r_n \ge z_n$ and is inverse S-shaped on $[-1, z_n]$ with inflection point $\ell_n \le z_n$.

These three cases together imply that there exists an $N \ge 0$ such that for all $n \ge N$ it holds that $W_n(\cdot)$ is strictly S-shaped on $[z_n, 1]$ with inflection point $r_n \in [z_n, 1]$ while strictly inverse S-shaped on $[-1, z_n]$ with inflection point $\ell_n \in [-1, z_n]$. Moreover, in the limit we have

$$\lim_{n\to\infty} \ell_n = \ell^* \quad \text{and} \quad \lim_{n\to\infty} r_n = r^* \ . \tag{C.11}$$

Finally, we show that $N = 0$ if $g(\cdot)$ is log-concave and ρ is sufficiently close to zero. By (C.7) we have

$$W_n''(\theta) = \hat{g}_n(\theta;q)\left(\theta - \phi_n(\theta)\right)\left\{\frac{2-\phi_n'(\theta)}{\theta - \phi_n(\theta)} + \frac{\hat{g}_n'(\theta;q)}{\hat{g}_n(\theta;q)}\right\} \tag{C.12}$$

We only consider interval $[z_n, 1]$ and show that $W_n(\cdot)$ is strictly S-shaped on it for all $n \ge 0$.[8] Observe that $\theta - \phi_n(\theta) > 0$ holds for all $\theta > z_n$ due to the single-crossing property. It therefore suffices to show that the overall term in the curly brackets in (C.12) is strictly decreasing. When $g(\cdot)$ is log-concave, it follows from Proposition A.1 in Appendix A that $\hat{g}_n(\cdot;q)$ is strictly log-concave, and hence $\frac{\hat{g}_n'(\theta;q)}{\hat{g}_n(\theta;q)}$ is strictly decreasing, for all $n \ge 0$. Next we show that $\frac{2-\phi_n'(\theta)}{\theta-\phi_n(\theta)}$ is also strictly decreasing in θ on $(z_n, 1]$ for all $n \ge 0$. Let $\xi_n(\theta) := \frac{2-\phi_n'(\theta)}{\theta-\phi_n(\theta)}$. Simple algebra reveals that $\xi_n'(\theta) < 0$ if and only if

$$\phi_n''(\theta)\left(\theta - \phi_n(\theta)\right) + \left(2 - \phi_n'(\theta)\right)\left(1 - \phi_n'(\theta)\right) \ge 0 \tag{C.13}$$

Recall from the definition of $\phi_n(\theta)$ (cf. (B.1)) that $\phi_n'(\theta) = \rho \sum_{j=1}^{n+1} w_j \cdot \varphi_j'(\theta;q,n)$ and $\phi_n''(\theta) = \rho \sum_{j=1}^{n+1} w_j \cdot \varphi_j''(\theta;q,n)$. As $\rho \to 0$ we have $\phi_n'(\theta) \to 0$ and $\phi_n''(\theta) \to 0$ uniformly for all $\theta \in [\underline{v}, \overline{v}]$. Therefore, the left-hand side of (C.13) converges to 2 for all $\theta \in [-1,1]$ as $\rho \to 0$. This implies $\xi_n'(\theta) < 0$ for all $\theta \in [-1,1]$ and $n \ge 0$ if ρ is sufficiently close to zero. This concludes the proof.

[8] The proof for the inverse S-shape property on $[-1, z_n]$ is analogous and hence omitted.

D Omitted Proofs for Section 6.1: Asymptotic Results

D.1 Proof of Lemma 6

To show part (1), recall that $\phi^* = \phi(v_q^*)$. Together with $\phi^* \in (-1,1)$ and $v_q^* = \phi^*$, this implies $v_q^* = \phi(v_q^*) \in (-1,1)$. (12) then implies $z^* = \phi^* = v_q^*$. To show part (2), note that $\phi(\cdot)$ is a constant and equal to χ when $\rho = 0$. Hence, by (12), $z^* = \phi^* = \chi$ in this case. If $\rho > 0$ then $\phi(\cdot)$ is a strictly increasing function. Together with $\phi^* = \phi(v_q^*)$, we obtain that $v^* > (<)\phi^*$ implies $z^* < (>)\phi^*$.

D.2 Proof of Theorem 2

To prove Theorem 2 we introduce two Lemmas D.1 and D.2, which respectively characterize the asymptotically optimal sender payoff and the set of censorship policies that generate this payoff.

Lemma D.1. *Suppose $v_q^*, \phi^* \in [-1,1]$ and let W_n be the value of problem (MP) for any $n \geq 0$. Then*

$$W^* := \lim_{n\to\infty} W_n = \begin{cases} \int_{\underline{\phi}}^1 (k-\phi^*)dF(k), & \textit{if } \phi^* < \underline{\phi} \\ \int_{\phi^*}^1 (k-\phi^*)dF(k), & \textit{if } \underline{\phi} \leq \phi^* \leq \overline{\phi} \\ \int_{\overline{\phi}}^1 (k-\phi^*)dF(k), & \textit{if } \phi^* > \overline{\phi} \end{cases} \tag{D.1}$$

Proof of Lemma D.1. We start by presenting the sender's asymptotic persuasion problem. Recall that $W_n(\cdot)$ is the sender's indirect utility function and let $W(\theta) := \lim_{n\to\infty} W_n(\theta)$. Then

$$W(\theta) = \begin{cases} \theta - \phi^*, & \text{if } \theta > v_q^* \\ 0, & \text{if } \theta < v_q^* \end{cases} .$$

This is because $v^{(nq+1)} \xrightarrow{p} v_q^*$ and $\varphi_n(v) \xrightarrow{a.s.} \phi^*$. If $\theta < v_q^*$, then $v^{(nq+1)} > \theta$ almost surely so that the status quo is maintained with probability one as $n \to \infty$. The sender's payoff thus converges to 0. Conversely, if $\theta > v_q^*$, then the reform is passed with probability one as $n \to \infty$ so that the sender's payoff converges to $\lim_{n\to\infty} (\theta - \varphi_n(v)) = \theta - \phi^*$. Let

$$\widetilde{W}(\theta) := \begin{cases} W(\theta), & \text{if } \theta \neq v_q^* \\ \max\{v_q^* - \phi^*, 0\}, & \text{if } \theta = v_q^* \end{cases} .$$

The sender's asymptotic persuasion problem is then

$$\max_{H \in \Delta([-1,1])} \int_{-1}^1 \widetilde{W}(\theta) dH(\theta), \text{ s.t. } F \succeq_{MPS} H \tag{D.2}$$

Because $\widetilde{W}(\cdot)$ is upper semi-continuous, problem (D.2) always admits a solution. We denote the value to (D.2) by W^* and characterize it using Theorem 1 of Dworczak and Martini (2019). For ease of reference we restate their theorem applied to our problem in the following observation.

Observation D.1. *(Theorem 1 of Dworczak and Martini (2019)) If there exists some $H \in \Delta([-1,1])$ and a convex function $p(\cdot)$ on $[-1,1]$ with $p(\cdot) \geq \widetilde{W}(\cdot)$ that satisfy*

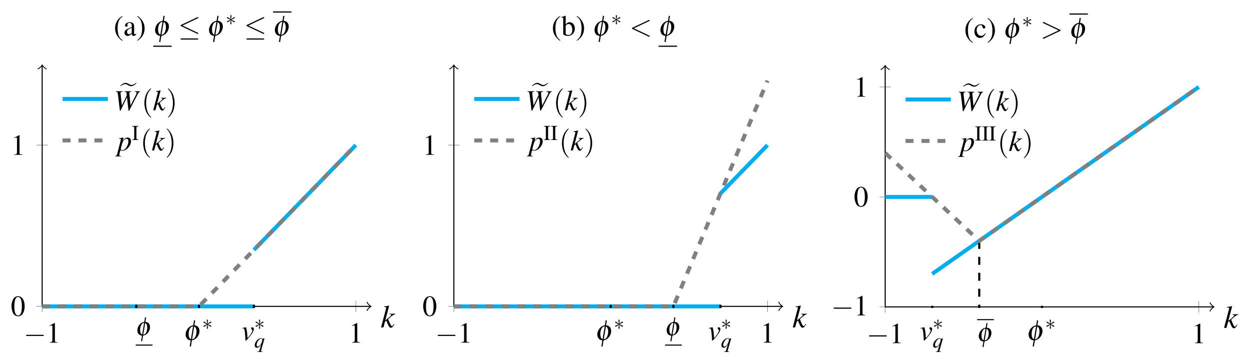
$$supp(H) \subseteq \left\{\theta : p(\theta) = \widetilde{W}(\theta)\right\}, \text{ and} \tag{D.3}$$

$$\int_{-1}^{1} p(\theta)dH(\theta) = \int_{-1}^{1} p(\theta)dH(\theta), \text{ and} \tag{D.4}$$

$$F \succeq_{MPS} H, \tag{D.5}$$

then H is a solution to (D.2) and $W^* = \int_{-1}^{1} p(\theta)dH(\theta)$.

Figure D.1: Illustration for the proof of Lemma D.1



Note: In all three panels, the blue solid lines denote $\widetilde{W}(\cdot)$ and the gray dashed lines denote the auxiliary functions $p(\cdot)$.

We distinguish between three cases.

Case 1: $\underline{\phi} \leq \phi^* \leq \overline{\phi}$. In this case we have $\mathbb{E}_F[k|k \geq \phi^*] \geq v_q^*$ and $\mathbb{E}_F[k|k \leq \phi^*] \leq v_q^*$. Let $p^{\mathrm{I}}(\theta) := \max\{\theta - \phi^*, 0\}$ for $\theta \in [-1,1]$. $p^{\mathrm{I}}(\cdot)$ is illustrated in Figure D.1a. Consider the cutoff policy $\mathscr{P}(\phi^*)$ and let $\underline{H}^{\mathrm{I}} = H_{\mathscr{P}(\phi^*)}$. The following conditions are easy to verify: (i) $p^{\mathrm{I}}(\cdot)$ is convex and $p^{\mathrm{I}}(\cdot) \geq \widetilde{W}(\cdot)$ on $[-1,1]$; (ii) $F \succeq_{MPS} \underline{H}^{\mathrm{I}}$ and $\int_{-1}^{1} p^{\mathrm{I}}(\theta)d\underline{H}^{\mathrm{I}}(\theta) = \int_{-1}^{1} p^{\mathrm{I}}(\theta)dF(\theta)$, and (iii)

$$\begin{aligned}
\operatorname{supp}(\underline{H}^{\mathrm{I}}) &= \{\mathbb{E}_F[k|k \leq \phi^*], \mathbb{E}_F[k|k \geq \phi^*]\} \\
&\subset \left\{\theta \middle| p^{\mathrm{I}}(\theta) = \widetilde{W}(\theta)\right\} = \begin{cases} [-1,1], & \text{if } v_q^* = \phi^* \\ [-1, v_q^*] \cup [\phi^*, 1], & \text{if } v_q^* < \phi^* \\ [-1, \phi^*] \cup [v_q^*, 1], & \text{if } v_q^* > \phi^* \end{cases}
\end{aligned}$$

Therefore, by Observation D.1, $\underline{H}^{\mathrm{I}}$ solves problem (D.2) and the value of (D.2) equals

$$\int_{-1}^{1} p^{\mathrm{I}}(k)dF(k) = \int_{-1}^{1} \max\{k-\phi^*,0\}dF(k) = \int_{\phi^*}^{1}(k-\phi^*)dF(k) \tag{D.6}$$

Case 2: $\phi^* < \underline{\phi}$. For this case we use the definition of $\underline{\phi}$ (cf. (14)) and obtain $\underline{\phi} \in (-1,1)$ and $\mathbb{E}_F[k|k \geq \underline{\phi}] = v_q^*$. Let

$$p^{\mathrm{II}}(\theta) := \begin{cases} \frac{v_q^*-\phi^*}{v_q^*-\underline{\phi}}\left(\theta-\underline{\phi}\right), & \text{if } \underline{\phi} \leq \theta \leq 1 \\ 0, & \text{if } -1 \leq \theta < \underline{\phi} \end{cases}.$$

$p^{\mathrm{II}}(\cdot)$ is illustrated in Figure D.1b. Consider cutoff policy $\mathscr{P}(\underline{\phi})$ and let $\underline{H}^{\mathrm{II}} = H_{\mathscr{P}(\underline{\phi})}$. The following conditions are easy to verify: (i) $p^{\mathrm{II}}(\cdot)$ is convex and $p^{\mathrm{II}}(\cdot) \geq \widetilde{W}(\cdot)$ on $[-1,1]$; (ii) $F \succeq_{MPS} \underline{H}^{\mathrm{II}}$ and $\int_{-1}^{1} p^{\mathrm{II}}(\theta)d\underline{H}^{\mathrm{II}}(\theta) = \int_{-1}^{1} p^{\mathrm{II}}(\theta)dF(\theta)$; and (iii)

$$\mathrm{supp}(\underline{H}^{\mathrm{II}}) = \left\{\mathbb{E}_F[k|k<\underline{\phi}], v_q^*\right\} \subset \left\{\theta \middle| p^{\mathrm{II}}(\theta) = \widetilde{W}(\theta)\right\} = [-1,\underline{\phi}] \cup \{v_q^*\}.$$

Hence, $\underline{H}^{\mathrm{II}}$ solves problem (D.2) and the value of (D.2) equals

$$\begin{aligned}\int_{-1}^{1} p^{\mathrm{II}}(k)dF(k) &= \left(1-F(\underline{\phi})\right)\frac{v_q^*-\phi^*}{v_q^*-\underline{\phi}}\left(\mathbb{E}_F[k|k>\underline{\phi}]-\underline{\phi}\right) \\ &= \left(1-F(\underline{\phi})\right)\left(v_q^*-\phi^*\right) = \int_{\underline{\phi}}^{1}(k-\phi^*)\,dF(k)\end{aligned} \tag{D.7}$$

Case 3: $\phi^* > \overline{\phi}$. For this case we use the definition of $\overline{\phi}$ (cf. (13)) and obtain $\overline{\phi} \in (-1,1)$ and $\mathbb{E}_F[k|k \leq \overline{\phi}] = v_q^*$. Let

$$p^{\mathrm{III}}(\theta) = \begin{cases} \theta-\phi^*, & \text{if } \overline{\phi} \leq 1 \\ \frac{\overline{\phi}-\phi^*}{\overline{\phi}-v_q^*}\left(\theta-v_q^*\right), & \text{if } -1 \leq \theta < \overline{\phi} \end{cases}.$$

$p^{\mathrm{III}}(\cdot)$ is illustrated in Figure D.1c. Consider cutoff policy $\mathscr{P}(\overline{\phi})$ and let $\underline{H}^{\mathrm{III}} = H_{\mathscr{P}(\overline{\phi})}$. The following conditions are again easy to verify: (i) $p^{\mathrm{III}}(\cdot)$ is convex and $p^{\mathrm{III}}(\cdot) \geq \widetilde{W}(\cdot)$; (ii) $F \succeq_{MPS} \underline{H}^{\mathrm{III}}$ and $\int_{-1}^{1} p^{\mathrm{III}}(\theta)d\underline{H}^{\mathrm{III}}(\theta) = \int_{-1}^{1} p^{\mathrm{III}}(\theta)dF(\theta)$; and (iii)

$$\mathrm{supp}(\underline{H}^{\mathrm{III}}) = \left\{\mathbb{E}_F\left[k|k>\overline{\phi}\right], v_q^*\right\} \subset \left\{\theta \middle| p^{\mathrm{III}}(\theta) = \widetilde{W}(\theta)\right\} = \{v_q^*\} \cup [\overline{\phi},1]$$

Following analogous arguments as in previous cases, we can establish that $\underline{H}^{\mathrm{III}}$ solves problem

(D.2) and the value of (D.2) equals

$$\begin{aligned}\int_{-1}^{1} p^{\mathrm{III}}(k)dF(k) &= F\left(\overline{\phi}\right)\frac{\overline{\phi}-\phi^*}{\overline{\phi}-v_q^*}\left(\mathbb{E}[k|k<\overline{\phi}]-v_q^*\right)+\int_{\overline{\phi}}^{1}(k-\phi^*)\,dF(k)\\ &= \int_{\phi}^{1}(k-\phi^*)dF(\theta)\end{aligned} \tag{D.8}$$

Taken together, (D.6) to (D.8) and Observation D.1 imply (D.1). □

Lemma D.2. *Suppose $v_q^*, \phi^* \in [-1,1]$ and let $\mathscr{P}^*$ denote the set of censorship policies that generate the asymptotically optimal payoff W^* given by* (D.1)*. Then $\mathscr{P}^*$ is characterized as follows.*

1. *If $\phi^* = v_q^*$ then $\mathscr{P}^* = \{\mathscr{P}(a,b):\ -1 \le a \le \phi^* \le b \le 1\}$.*
2. *If $\phi^* < v_q^*$ then $\mathscr{P}^* = \{\mathscr{P}(a,b):\ -1 \le a \le b = \min\{\phi^*, \overline{\phi}\}\}$.*
3. *If $\phi^* > v_q^*$ then $\mathscr{P}^* = \left\{\mathscr{P}(a,b):\ \max\{\phi^*, \underline{\phi}\} = a \le b \le 1\right\}$.*

Proof of Lemma D.2. Observe that (D.1) can be rewritten as $W^* = \int_{t^*}^{1}(k-\phi^*)dF(k)$ where $t^* = \text{median}\left\{\overline{\phi}, \phi^*, \underline{\phi}\right\}$. We distinguish between three cases:

Case 1. $\overline{\phi} \ge \phi^* \ge \underline{\phi}$ and therefore $W^* = \int_{\phi^*}^{1}(k-\phi^*)dF(k)$. If $v_q^* = \phi^*$, then all $\mathscr{P}(a,b)$ with $a \le \phi^* \le b$ yield the same optimal asymptotic payoff W^*. If $\phi^* < v_q^*$, then $a \le b = \phi^*$ is necessary. If $\phi^* > v_q^*$ then $\phi^* = a \le b$ is necessary.

Case 2: $\phi^* > \overline{\phi}$ and therefore $W^* = \int_{\overline{\phi}}^{1}(k-\phi^*)dF(k)$. In this case it must hold that $\phi^* > v_q^*$. A censorship policy $\mathscr{P}(a,b)$ can implement the same outcome if and only if $\overline{\phi} = a \le b$.

Case 3: $\phi^* < \underline{\phi}$ and therefore $W^* = \int_{\underline{\phi}}^{1}(k-\phi^*)dF(k)$. In this case it must hold that $\phi^* < v_q^*$. A censorship policy $\mathscr{P}(a,b)$ can implement the same outcome if and only if $a \ge b = \underline{\phi}$.

These together complete the proof of Lemma D.2. □

Now we prove Theorem 2.

Proof of Theorem 2. Consider any pair of sequences $\{b_n\}_{n\ge N}$ and $\{a_n\}_{n\ge N}$ of optimal thresholds. Because both sequences are bounded on a closed interval, by the Bolzano–Weierstrass Theorem they must contain at least one convergent subsequence each. Let b^* and a^* denote the limits of these convergent subsequences. In what follows we shall explicitly characterize b^* and a^*, and then show that all sub-sequences of $\{a_n\}_{n\ge N}$ and $\{b_n\}_{n\ge N}$ converge to them so that $\{a_n\}_{n\ge N}$ and $\{b_n\}_{n\ge N}$ indeed converge.[9]

[9] Here we exploit the following observation: let $\{x_n\}$ be a sequence on a bounded closed interval and suppose all its convergent subsequences have the same limit x^*; then x_n converges to x^*. To see this, suppose instead that x_n does not converge to x^*. Then there exists some $\varepsilon > 0$ and a subsequence $\left\{x_{n_j}\right\}$ indexed by $j = 1, 2, \cdots$ such that $\left|x_{n_j} - x^*\right| > \varepsilon$ holds for all n_j. Since x_{n_j} is bounded in a closed interval, by the Bolzano-Weierstrass Theorem it must contain a convergent subsequence. Yet this subsequence does not converge to x^*, leading to a contradiction.

On the one hand, asymptotic optimality requires that $\mathscr{P}(a^*,b^*) \in \mathscr{P}^*$ must hold, where $\mathscr{P}^*$ is characterized in Lemma D.2. On the other hand, by Lemma 5, the single-crossing property implies for sufficiently large n that there exists $z_n \in [-1,1]$ and $-1 \leq \ell_n \leq z_n \leq r_n \leq 1$ such that the following conditions must hold:

$$\ell_n \leq a_n \leq z_n \leq b_n \leq r_n$$

and $z_n \to z^*$, $\ell_n \to \min\{z^*, v_q^*\}$ and $r_n \to \max\{z^*, v_q^*\}$ for $v_q^* \in (-1,1)$.[10] We now distinguish between three cases. For all these cases recall that $\phi(v_q^*) = \phi^*$ and $\phi(\cdot)$ is non-decreasing.

Case 1. $\phi^* = v_q^*$. In this case $z^* = v_q^*$ because $\phi(v_q^*) = \phi^* = v_q^*$. Therefore, both ℓ_n and r_n converge to ϕ^*. By the squeeze theorem both a_n and b_n must also converge to ϕ^* and thus $a^* = b^* = \phi^*$.

Case 2. $\phi^* < v_q^*$. In this case $z^* \leq \phi^* < v_q^*$ so that $\ell_n \to z^*$ and hence $a^* = z^* < v_q^*$. Moreover, by part (2) of Lemma D.2, $b^* = \phi^*$if $\phi^* \in [\underline{\phi}, v_q^*)$ and $b^* = \underline{\phi}$ if $\phi^* < \underline{\phi}$.

Case 3. $\phi^* > v_q^*$. In this case $z^* \geq \phi^* > v_q^*$ so that $r_n \to z^*$ and hence $b^* = z^* > v_q^*$. Moreover, by part (3) of Lemma D.2 we have $a^* = \phi^*$if $\phi^* \in (v_q^*, \overline{\phi}]$ and $a^* = \overline{\phi}$ if $\phi^* > \overline{\phi}$.

These complete the characterizations of $a^*$and b^* and these apply for any convergent subsequences of a_n and b_n. Therefore, the limits of $\{a_n\}$ and $\{b_n\}$ exist and are equal to a^* and b^*, respectively. □

D.3 Proof of Proposition 1

It is straightforward from the definitions of $\overline{W}$ and W^{Full} that

$$\overline{W} = \int_{\phi^*}^{1} (k - \phi^*) f(k) dk \tag{D.9}$$

$$W^{\text{Full}} = \int_{v_q^*}^{1} (k - \phi^*) f(k) dk \tag{D.10}$$

This is because, as $n \to \infty$, in the omniscient scenario the sender prefers reform (status quo) if $k > (<) \phi^*$, while under full information the pivotal voter prefers reform (status quo) if $k > (<) v_q^*$.

Let $\gamma(x) := \int_x^1 (k - \phi^*) dF(k)$ for $x \in [-1,1]$. By (D.9) and (D.10) we have $\overline{W} = \gamma(\phi^*)$ and $W^{\text{Full}} = \gamma(v_q^*)$. Note that $\gamma'(x) = (\phi^* - x) f(x) > (<) 0$ for $x < (>) \phi^*$. This implies that $\gamma(x)$ is strictly increasing on $[-1, \phi^*)$ and strictly decreasing on $(\phi^*, 1]$. Therefore, $\overline{W} \geq W^{\text{Full}}$ and equality holds if and only if $\phi^* = v_q^*$ whenever $v_q^* \in (-1,1)$. Moreover, by Lemma D.1, $W^* = \gamma(\phi^*) = \overline{W}$ if $\underline{\phi} \leq \phi^* \leq \overline{\phi}$. These together establish statements (1) and (2). To show statement (3), consider $\phi^* > \overline{\phi}$ first. In this case, it follows from Lemma D.1 that $W^* = \gamma(\overline{\phi}$ with $\overline{\phi} \in (v_q^*, \phi^*)$. Since $\gamma(x)$ is strictly increasing for $x < \phi^*$, it holds that $\gamma(\phi^*) > \gamma(\overline{\phi}) > \gamma(v_q^*)$, or equivalently $\overline{W} > W^* > W^{\text{Full}}$. The proof for the case $\phi^* < \underline{\phi}$ is analogous.

[10] The limiting results for ℓ_n and r_n follow from (C.10) and (C.11) in Appendix C.2.

E Omitted Proofs for Section 6.2: Comparative Statics

In this appendix we prove the comparative static results presented in Section 6.2. We use two different approaches to establish these results.

E.1 First-order approach and proofs of Propositions 2 and 3

In this subsection we use the first-order approach to prove Propositions 2 and 3. We only prove the statements concerning b_n; the proofs for claims concerning a_n are similar and thus omitted.

Recall from (FOC: b_n) in Section 5.2 that the optimality condition for b_n is given by

$$\left(\tilde{b}_n - b_n\right) W_n'(\tilde{b}_n) \leq W_n(\tilde{b}_n) - W_n(b_n) \tag{E.1}$$

where $\tilde{b}_n = \mathbb{E}_F[k|k \geq b_n]$ and this condition is binding whenever $b_n \in (-1,1)$. Recall from (9) that $W_n(\theta) = \int_{\underline{v}}^{\theta} (\theta - \phi_n(x))\, \hat{g}_n(x;q)dx$ and its derivative is given by

$$W_n{}'(\theta) = \hat{G}_n(\theta;q) + (\theta - \phi_n(\theta))\, \hat{g}_n(\theta;q)$$

With these we have

$$\begin{aligned}
W_n(\tilde{b}_n) - W_n(b_n) &= \int_{\underline{v}}^{\tilde{b}_n} \left(\tilde{b}_n - \phi_n(x)\right) \hat{g}_n(x;q)dx - \int_{\underline{v}}^{b_n} (b_n - \phi_n(x))\, \hat{g}_n(x;q)dx \\
&= \int_{b_n}^{\tilde{b}_n} \left(\tilde{b}_n - \phi_n(x)\right) \hat{g}_n(x;q)dx + (\tilde{b}_n - b_n) \int_{\underline{v}}^{b_n} \hat{g}_n(x;q)dx \\
\left(\tilde{b}_n - b_n\right) W_n'(\tilde{b}_n) &= \left(\tilde{b}_n - b_n\right) \left[\left(\tilde{b}_n - \phi_n^m(\tilde{b}_n)\right) \hat{g}_n(\tilde{b}_n;q) + \int_{\underline{v}}^{\tilde{b}_n} \hat{g}_n(x;q)dx \right]
\end{aligned}$$

Plugging these into (E.1), we obtain that (E.1) is equivalent to

$$\tilde{b}_n - b_n \leq \int_{b_n}^{\tilde{b}_n} \frac{b_n - \phi_n(x)}{\tilde{b}_n - \phi_n(\tilde{b}_n)} \frac{\hat{g}_n(x;q)}{\hat{g}_n(\tilde{b}_n;q)} dx \tag{E.2}$$

By Lemma 5, for sufficiently large n it holds that $W_n(\cdot)$ is strictly S-shaped on $[z_n, 1]$ with some inflection point $r_n \in [z_n, 1]$. This implies that

$$\left(\mathbb{E}_F\left[k|k \geq x\right] - x\right) W_n'\left(\mathbb{E}_F\left[k|k \geq x\right]\right) - \left[W_n\left(\mathbb{E}_F\left[k|k \geq x\right]\right) - W_n\left(x\right)\right]$$

can cross zero at most once and from above as x increases from z_n to 1. In particular, suppose b_n satisfies (E.2) with equality and hold it fixed, then if any parameter change increases the value of the right-hand side of (E.2), then $\left(\tilde{b}_n - b_n\right) W_n'(\tilde{b}_n) - \left[W_n(\tilde{b}_n) - W_n(b_n)\right]$ will be negative following this parameter change. b_n must therefore decrease to restore equality. Comparative static analyses thus can done with the right-hand side of (E.2) alone. With this we can prove Propositions 2 and 3.

Proof of Proposition 2. Let $\gamma_n(x) := \sum_{j=1}^{n+1} w_j \varphi_j(x;q,n)$. If $\rho < 1$, it follows from (3) that

$$\phi_n(\cdot) = \rho \gamma_n(x) + (1-\rho)\chi$$

As is explained in the proof of Lemma 3 in Appendix B, under either condition (i) or (ii) of Lemma 3 it holds that $1-\phi'(x) > 0$ for all $x \in [\underline{v}, \overline{v}]$. Since $\phi_n(\cdot)$ converges uniformly to $\phi'(\cdot)$ (cf. Lemma 2), $1-\phi_n'(\cdot) > 0$ on $[\underline{v}, \overline{v}]$ must hold for sufficiently large n. This implies that $x - \phi_n(x)$ is strictly increasing. Moreover, because $\tilde{b}_n = \mathbb{E}_F[k|k \geq b_n] > b_n$ and $\phi_n(x)$ is non-decreasing (cf. Proposition B.2), for all $x \in (b_n, \tilde{b}_n)$ we have

$$1 > \frac{b_n - \phi_n(b_n)}{\tilde{b}_n - \phi_n(\tilde{b}_n)} \geq \frac{b_n - \phi_n(x)}{\tilde{b}_n - \phi_n(\tilde{b}_n)} = \frac{b_n - \rho\gamma_n(x) - (1-\rho)\chi}{\tilde{b}_n - \rho\gamma_n(\tilde{b}_n) - (1-\rho)\chi}$$

Consider any $\chi_{\text{I}} > \chi_{\text{II}}$. Observe that a decrease of χ from χ_{I} to χ_{II} induces a common increase on both the nominator and the denominator of $\frac{b_n - \phi_n(x)}{\tilde{b}_n - \phi_n(\tilde{b}_n)}$, which is smaller than one for all $x \in \left(b_n, \tilde{b}_n\right)$. This shift of χ therefore strictly increases the value of $\frac{b_n - \phi_n(x)}{\tilde{b}_n - \phi_n(\tilde{b}_n)}$ for all $x \in \left(b_n, \tilde{b}_n\right)$.[11] On the other hand, the term $\frac{\hat{g}_n(x;q)}{\hat{g}_n(\tilde{b}_n;q)}$ is independent of χ for all $x \in \left(b_n, \tilde{b}_n\right)$. These together implies that a shift of χ from χ_{I} to χ_{II} strictly increases the right-hand side of (E.2). Therefore, if $b_n \in (-1,1)$ under χ_{I} so that (E.2) is binding, such a shift of χ will make the right-hand side of (E.2) strictly higher than the left-hand side. b_n must strictly decrease to make (E.2) binding again or drop to -1. If $b_n = -1$ under χ_{I} so that (E.2) holds with '$\leq$', then this must remain to be the case after the shift of χ so that the optimal b_n remains to be -1. These together show that b_n is non-increasing as χ decreases and thus prove the claim for b_n in Proposition 2. □

Next we prove Proposition 3. To do so we introduce an auxiliary result.

Lemma E.1. *For any $0 < y < z < 1$, $\int_y^z \frac{\tau_n(x;q)}{\tau_n(z;q)} dx$ is strictly decreasing in q, where $\tau_n(x;q)$ is defined by (A.1) in Appendix A.*[12]

Proof of Lemma E.1. For any pair of $(x,y) \in (0,1)^2$ and $q \in (0,1)$, define

$$\Delta\psi(x,y;q) := q \ln\frac{x}{y} + (1-q)\ln\frac{1-x}{1-y} = \ln\frac{1-x}{1-y} + q\left(\ln\frac{x}{1-x} - \ln\frac{y}{1-y}\right) \qquad \text{(E.3)}$$

[11] This follows from the fact that $\frac{a+c}{b+c} > \frac{a}{b}$ for all $b,c > 0$ and $b > a$.

[12] Here we implicitly assume nq is an integer for ease of exposure. If this is not the case, then just replace q with $\hat{q} = \lfloor nq \rfloor / n$ and the all arguments hold for $\hat{q}$.

It then follows from the definition of $\tau_n(\cdot;q)$ that

$$\ln\frac{\tau_n(x;q)}{\tau_n(y;q)} = n\left(q\ln\frac{x}{y}+(1-q)\ln\frac{1-x}{1-y}\right) = n\Delta\psi(x,y;q)$$

We can thus rewrite $\int_y^z \frac{\tau_n(x;q)}{\tau_n(z;q)}dx$ as

$$\int_y^z \frac{\tau_n(x;q)}{\tau_n(z;q)}dx = \int_y^z e^{n\Delta\psi(x,z;q)}dx$$

Using (E.3) and the fact that $\ln\frac{x}{1-x}$ is strictly increasing in x, we obtain for all $y<z$ that

$$\frac{\partial}{\partial q}\int_y^z \frac{\tau_n(x;q)}{\tau_n(z;q)}dx = n\int_y^z e^{n\Delta\psi(x,z;q)}\left(\ln\frac{x}{1-x}-\ln\frac{z}{1-z}\right)dx<0$$

This implies the strict decreasing property stated in this lemma. □

Proof of Proposition 3. Recall from (A.5) that $\hat{g}_n(x;q)=\tau_n\left(G(x);q\right)g(x)$ for all $x\in[\underline{v},\overline{v}]$. Plugging this into (E.2), we obtain

$$\begin{aligned}\int_{b_n}^{\tilde{b}_n}\frac{b_n-\phi_n(x)}{\tilde{b}_n-\phi_n(\tilde{b}_n)}\frac{\hat{g}_n(x;q)}{\hat{g}_n(\tilde{b}_n;q)}dx &= \int_{b_n}^{\tilde{b}_n}\frac{b_n-\phi_n(x)}{\tilde{b}_n-\phi_n(\tilde{b}_n)}\frac{\tau_n\left(G(x);q\right)g(x)}{\tau_n\left(G(\tilde{b}_n);q\right)g(\tilde{b}_n)}dx \\ &= \frac{1}{g(\tilde{b}_n)}\int_{G(b_n)}^{G(\tilde{b}_n)}\frac{b_n-\phi_n\left(G^{-1}(y)\right)}{\tilde{b}_n-\phi_n(\tilde{b}_n)}\frac{\tau_n\left(y;q\right)}{\tau_n\left(G(\tilde{b}_n);q\right)}dy\end{aligned} \tag{E.4}$$

For $\rho=0$ we have $\phi_n(x)=\chi$ for all $x\in[\underline{v},\overline{v}]$ and therefore $W_n(\theta)=(\theta-\chi)\hat{G}_n(\theta;q)$ (cf. (3) and (9)). Plugging $W_n(\theta)=(\theta-\chi)\hat{G}_n(\theta;q)$ into (E.4), we obtain

$$\int_{b_n}^{\tilde{b}_n}\frac{b_n-\phi_n(x)}{\tilde{b}_n-\phi_n(\tilde{b}_n)}\frac{\hat{g}_n(x;q)}{\hat{g}_n(\tilde{b}_n;q)}dx = \frac{1}{g(\tilde{b}_n)}\left(\frac{b_n-\chi}{\tilde{b}_n-\chi}\right)\int_{G(b_n)}^{G(\tilde{b}_n)}\frac{\tau_n(x)}{\tau_n(G(\tilde{b}_n)}dx \tag{E.5}$$

Because $\tilde{b}_n>b_n>\chi$ and $G(\tilde{b}_n)>G(b_n)$, Lemma E.1 implies that (E.5) is strictly decreasing in q. Therefore, the right-hand side of (E.2) strictly increases as q rises from q_{I} to q_{II} for all $q_{\mathrm{I}}<q_{\mathrm{II}}$. If $b_n\in(-1,1)$ under q_{I} so that (E.2) is binding, then such a shift of q will make (E.2) hold with '$>$' so that b_n must strictly increase to regain equality or up to 1. If $b_n=-1$ under q_{I}, then (E.2) holds with '$\leq$'. The shift of q either (i) retains (E.2) with '$\leq$' so that the optimal b_n is still -1, or it shifts '$\leq$' to '$>$' so that the optimal $b_n>-1$. These together show that b_n is non-decreasing as q increases and thus prove the claim for b_n in Proposition 3. □

E.2 Limiting approach for comparative statics

In this subsection we use the limiting approach to prove Proposition 4 and establish Proposition E.1 below, which is an analog of Proposition 2 for the welfare weighting function $w(\cdot)$ when $\rho > 0$.

Suppose the single-crossing property holds and let $a^* := \lim_{n\to\infty} a_n$ and $b^* := \lim_{n\to\infty} b_n$ be the limits of optimal thresholds characterized in Theorem 2. Our comparative statics primarily concern how a^* and b^* vary with voting rule q and a pro-social sender's welfare weighting function $w(\cdot)$. By Theorem 2, this boils down to understanding how these factors affect ϕ^*, v_q^* and z^*. For ϕ^* these are already summarized in Lemma B.5 in Appendix B. From the definition of v_q^* it is obvious that it is strictly increasing in q and independent of $w(\cdot)$. As for $z*$, Lemma E.2 below shows that when both G and $1-G$ are strictly log-concave it inherits the comparative statics of ϕ^*.

Lemma E.2. *Suppose both G and $1-G$ are strictly log-concave. Then (i) z^* weakly decreases as $w(\cdot)$ shifts from $w^I(\cdot)$ to $w^{II}(\cdot)$, where $w^I(\cdot) \succeq_{FOSD} w^{II}(\cdot)$; (ii) z^* is weakly decreasing in q.*

Proof of Lemma E.2. By Lemma 3, strict log-concavities of G and $1-G$ ensure the single-crossing property and hence the existence of a unique z^* for all $w(\cdot)$ and q. The definition of z^* (cf. (12)) implies that it must decrease if function $\phi(\cdot)$ systematically shifts downward – i.e., $\phi(x)$ strictly decreases for all $x \in (\underline{v}, \overline{v})$ – after some shift of $w(\cdot)$ or q. The decreasing properties (i) and (ii) then follow from Lemma B.4, which claims that $\phi(\cdot)$ shifts downwards if q increases or if $w(\cdot)$ varies from some $w^{\mathrm{I}}(\cdot)$ to $w^{\mathrm{II}}(\cdot)$ with $w^{\mathrm{I}}(\cdot) \succeq_{FOSD} w^{\mathrm{II}}(\cdot)$, when $\rho > 0$. $\square$

With these we are ready to prove Proposition 4 and establish Proposition E.1. For the proof of Proposition 4 we will frequently use the fact that the two boundaries $\overline{\phi}$ and $\underline{\phi}$ defined by equations (13) and (14) in the main text are both non-decreasing functions of v_q^*. We therefore write them as $\overline{\phi}(v_q^*)$ and $\underline{\phi}(v_q^*)$ in the proof to make this dependence explicit.

Proof of Proposition 4. We prove Proposition 4 by construction. For any pair of q_{I} and q_{II}, we use a_i^* and b_i^* to denote the thresholds of the asymptotically optimal censorship policy under $q = q_i$ for $i \in \{\mathrm{I}, \mathrm{II}\}$. We assume $\phi^* \in (-1, 1)$ and let $\widehat{q} := G(\phi^*)$; under $q = \widehat{q}$ we have $\phi^* = G^{-1}(q) = v_q^*$.

First, suppose q_{I} and q_{II} satisfy (i) $\widehat{q} \le q_{\mathrm{I}} < q_{\mathrm{II}}$, (ii) $\phi^* \in \left[\underline{\phi}(v_{q_{\mathrm{I}}}^*), \overline{\phi}(v_{q_{\mathrm{I}}}^*)\right]$, and (iii) $\phi^* < \underline{\phi}(v_{q_{\mathrm{II}}}^*)$. Then, by Theorem 2, $(a_{\mathrm{I}}^*, b_{\mathrm{I}}^*) = (z_{\mathrm{I}}^*, \phi^*)$ and $(a_{\mathrm{II}}^*, b_{\mathrm{II}}^*) = \left(z_{\mathrm{II}}^*, \underline{\phi}(v_{q_{\mathrm{II}}}^*)\right)$. Since $\underline{\phi}(v_{q_{\mathrm{II}}}^*) > \phi^*$ and $z_{\mathrm{II}}^* < z_{\mathrm{I}}^*$ (by Lemma E.2), we get $a_{\mathrm{II}}^* < a_{\mathrm{I}}^* \le b_{\mathrm{I}}^* < b_{\mathrm{II}}^*$. This implies for sufficiently large n that a_n decreases and b_n increases as q shifts from q_{I} to q_{II}. This shows case 1 of Proposition 4 is possible.

To show that case 2 is also possible, consider any q_{I} and q_{II} that satisfy (i) $q_{\mathrm{I}} < q_{\mathrm{II}} \le \widehat{q}$, (ii) $\phi^* > \overline{\phi}(v_{q_{\mathrm{I}}}^*)$, and (iii) $\phi^* \in \left[\underline{\phi}(v_{q_{\mathrm{II}}}^*), \overline{\phi}(v_{q_{\mathrm{II}}}^*)\right]$. By Theorem 2, $(a_{\mathrm{I}}^*, b_{\mathrm{I}}^*) = \left(\overline{\phi}(v_{q_{\mathrm{I}}}^*), z_{\mathrm{I}}^*\right)$ and $(a_{\mathrm{II}}^*, b_{\mathrm{II}}^*) = (\phi^*, z_{\mathrm{II}}^*)$. In this case we have $a_{\mathrm{I}}^* < a_{\mathrm{II}}^* \le b_{\mathrm{II}}^* < b_{\mathrm{I}}^*$. So, for sufficiently large n, a_n increases while b_n decreases as q varies from q_{I} to q_{II}.

To show that case 3 is also possible, consider any q_{I} and q_{II} that satisfy (i) $q_{\text{I}} < \widehat{q} < q_{\text{II}}$, (ii) $\phi^* \in \left[\underline{\phi}(v^*_{q_{\text{I}}}), \overline{\phi}(v^*_{q_{\text{I}}})\right]$, and (iii) $\phi^* \in \left[\underline{\phi}(v^*_{q_{\text{II}}}), \overline{\phi}(v^*_{q_{\text{II}}})\right]$. By Theorem 2, $(a^*_{\text{I}}, b^*_{\text{I}}) = (\phi^*, z^*_{\text{I}})$ and $(a^*_{\text{II}}, b^*_{\text{II}}) = (z^*_{\text{II}}, \phi^*)$. In this case we have $a^*_{\text{I}} > a^*_{\text{II}}$ and $b^*_{\text{I}} > b^*_{\text{II}}$. So, for sufficiently large n, both a_n and b_n decrease as q varies from q_{I} to q_{II}.

Finally, to show that case 4 is also possible, consider any q_{I} and q_{II} that satisfy (i) $q_{\text{I}} < \widehat{q} < q_{\text{II}}$, (ii) $\overline{\phi}(v^*_{q_{\text{I}}}) < \phi^* < \underline{\phi}(v^*_{q_{\text{II}}})$. Then, by Theorem 2, we have $(a^*_{\text{I}}, b^*_{\text{I}}) = \left(\overline{\phi}(v^*_{q_{\text{I}}}), z^*_{\text{I}}\right)$ and $(a^*_{\text{II}}, b^*_{\text{II}}) = \left(z^*_{\text{II}}, \underline{\phi}(v^*_{q_{\text{II}}})\right)$. By Lemma E.2, $z^*_{q_I} < \phi^* < z^*_{q_{II}}$ must hold for all $\rho > 0$. If, however, $\rho \to 0$, then $\phi(\cdot)$ must be close to a flat line so that both $z^*_{q_I}$ and $z^*_{q_{II}}$ shall be arbitrarily close to ϕ^*. This implies that $\overline{\phi}\left(v^*_{q_I}\right) < z^*_{q_{II}}$ and $z^*_{q_I} < \underline{\phi}\left(v^*_{q_{II}}\right)$ must hold for ρ sufficiently close to 0. In such case we indeed have $a^*_I < a^*_{II}$ and $b^*_I < b^*_{II}$. □

Proposition E.1. *Suppose $\rho > 0$ and $v^*_q \in (-1,1)$. Let $w^I(\cdot)$ and $w^{II}(\cdot)$ be two absolutely continuous cdfs on $[-1,1]$ that satisfy (i) $w^I(\cdot) \succeq_{FOSD} w^{II}(\cdot)$, and (ii) $\phi^*_i \in \left(\underline{\phi}(v^*_q), \overline{\phi}(v^*_q)\right)$ holds under $w^i(\cdot)$ for at least one $i \in \{I, II\}$. Then, for sufficiently large n, both a_n and b_n decrease as the weighting function $w(\cdot)$ shifts from $w^I(\cdot)$ to $w^{II}(\cdot)$.*

Proof of Proposition E.1. We use a^*_i and b^*_i to denote the thresholds of the asymptotically optimal censorship policy under $w(\cdot) = w^i(\cdot)$ for $i \in \{\text{I}, \text{II}\}$. For ease of exposure we focus on the case where $\phi^*_i \in \left(\underline{\phi}(v^*_q), \overline{\phi}(v^*_q)\right)$ holds under $w^i(\cdot)$ for both $i \in \{\text{I}, \text{II}\}$.[13] In this case, Theorem 2 implies $a^*_i = \min\{z^*_i, \phi^*_i\}$ and $b^*_i = \max\{z^*_i, \phi^*_i\}$ for both $i \in \{\text{I}, \text{II}\}$. By Lemma B.5, we have $\phi^*_{\text{I}} > \phi^*_{\text{II}}$ and $z^*_{\text{I}} > z^*_{\text{II}}$ (equality holds only if both values are -1). These together imply (i) $b^*_{\text{I}} > b^*_{\text{II}}$ and (ii) $a^*_{\text{I}} \geq a^*_{\text{II}}$ (equality holds only if both values are -1). Hence, for sufficiently large n, both a_n and b_n decrease as q shifts from $w^{\text{I}}(\cdot)$ to $w^{\text{II}}(\cdot)$. □

F Omitted Equilibrium Derivations and Proofs for Section 7

In this Appendix we solve for the equilibria of the competitive persuasion model in Section 7 and prove Theorems 3 and 4.

Equilibrium. When there are $|M| \geq 2$ senders, let $\pi = \left\langle \{\pi_m\}_{m \in M} \right\rangle$ be any joint information policy induced by all senders and let H_π denote the distribution of the posterior means induced by π. We say that H_π is *unimprovable for sender $m \in M$* if he has no incentive to reveal more information. For $m \in M$, let $\mathscr{H}_m$ denote the set of all unimprovable distributions. The set of distributions H that are *unimprovable for all senders* is then $\mathscr{H} = \cap_{m \in M} \mathscr{H}_m$. By Proposition 2 of Gentzkow and Kamenica (2017*b*), π can be sustained in equilibrium if and only if $H_\pi \in \mathscr{H}$.

[13] The proof is almost identical for the case where only one weighting function satisfies this condition.

To further solve for the unimprovable H's, we introduce some useful observations about properties of solutions to a general linear persuasion problem of the following kind:

$$\max_{H\in\Delta([\underline{\kappa},\overline{\kappa}])}\int_{\underline{\kappa}}^{\overline{\kappa}}U(\theta)dF(\theta),\quad \text{s.t. } F\big|_{[\underline{\kappa},\overline{\kappa}]}\succeq_{MPS} H \qquad \text{(MP')}$$

where $U(\cdot)$ is a sender's utility function defined on some closed interval $[\underline{\kappa},\overline{\kappa}]\subseteq[-1,1]$ and $F\big|_{[\underline{\kappa},\overline{\kappa}]}$ is the cdf of prior F truncated on interval $[\underline{\kappa},\overline{\kappa}]$.[14] We assume throughout this appendix that $U(\cdot)$ is twice continuously differentiable on $[-1,1]$.[15]

The first observation, due to Theorem 4 of Dworczak and Martini (2019), provides a convenient way to verify whether an induced distribution of posterior means H is unimprovable for a sender with utility function $U(\cdot)$ by solving a monopolistic persuasion problem with his utility function modified by its convex translations.

Observation F.1. *(Theorem 4 of Dworczak and Martini (2019)) If $H\in\Delta([\underline{\kappa},\overline{\kappa}])$ is unimprovable for a sender with utility function $U(\cdot)$ on $[\underline{\kappa},\overline{\kappa}]$,[16] then there exists a convex function $\omega(\cdot)$ on $[\underline{\kappa},\overline{\kappa}]$ such that H is a solution to (MP') with $U(\cdot)$ therein replaced by $\widehat{U}(\cdot)=U(\cdot)+\omega(\cdot)$.*

Using Observation F.1, we establish Lemma F.1, which implies that any unimprovable H must induce cutoff partitions at z_n^m for all $m\in M$. As a result, any $H\in\mathscr{H}$ must be unimprovable separately on each segment of interval $[-1,1]$ partitioned by the z_n^m's.

Lemma F.1. *Suppose that the single-crossing property holds for a sender $m\in M$ and $\phi_n^m(x)-x$ crosses zero only once and from above at $z_n^m\in(-1,1)$. Then $H\succeq_{MPS}H_{\mathscr{P}(z_n^m)}$ must hold for any H that is unimprovable for sender m on $[-1,1]$.*

Proof of Lemma F.1. By Observation F.1, H is unimprovable for sender m if and only if H is a solution to (MP) with utility function $W_n^m(\cdot)$ replaced by $\widehat{W}_n^m(\cdot)=W_n^m(\cdot)+\omega_m(\cdot)$ for some convex $\omega_m(\cdot)$. By Lemma 4, $W_n^m(\cdot)$ satisfies the increasing slope property at z_n^m. We show that $\widehat{W}_n^m(\cdot)$ must also satisfy this property and hence $H\succeq_{MPS}H_{\mathscr{P}(z_n^m)}$ follows Lemma 4. To see why, observe that

$$\frac{\widehat{W}_n^m(x)-\widehat{W}_n^m(z_n^m)}{x-z_n^m}=\frac{W_n^m(x)-W_n^m(z_n^m)}{x-z_n^m}+\frac{\omega_m(x)-\omega_m(z_n^m)}{x-z_n^m}$$

and the latter term is non-decreasing in x because $\omega_m(\cdot)$ is convex. $\widehat{W}_n^m(x)$ thus satisfies the increasing slope property whenever $W_n^m(\cdot)$ does. □

[14] That is, $F\big|_{[\underline{\kappa},\overline{\kappa}]}(k)=\frac{F(k)-F(\underline{\kappa})}{F(\overline{\kappa})-F(\underline{\kappa})}$ for $k\in[\underline{\kappa},\overline{\kappa}]$ and it equals 1 (resp. 0) for $k>\overline{\kappa}$ (resp. $k<\underline{\kappa}$).

[15] This ensures that $U(\cdot)$ is *regular*: $[-1,1]$ can be partitioned into finitely many intervals on which $U(\cdot)$ is either strictly concave, strictly convex, or affine. This regularity property is needed for Observations F.1 and F.4 below.

[16] Formally, H is unimprovable for a sender with utility function $U(\cdot)$ on $[\underline{\kappa},\overline{\kappa}]$ if $\int_{\underline{\kappa}}^{\overline{\kappa}}U(\theta)dH(\theta)\geq\int_{\underline{\kappa}}^{\overline{\kappa}}U(\theta)d\widetilde{H}(\theta)$ holds for all $\widetilde{H}\in\Delta([\underline{\kappa},\overline{\kappa}])$ that satisfies the feasibility constraint $F\big|_{[\underline{\kappa},\overline{\kappa}]}\succeq_{MPS}\widetilde{H}\succeq_{MPS}H$.

The next two observations establish, for the strictly S-shaped and inverse-S-shaped utility functions, the unimprovability and best-response properties of censorship policies under competitive persuasion.

Observation F.2. *Let $U(\cdot)$ be sender m's utility function and suppose that $U(\cdot)$ is strictly S-shaped on $[\underline{\kappa}, \overline{\kappa}]$ with inflection point r. The following properties hold:*

1. *(Kolotilin, Mylovanov and Zapechelnyuk, 2022) The unique solution H to problem (MP') is induced by an upper censorship policy $\mathscr{P}(\underline{\kappa}, b)$ with $b \geq \underline{\kappa}$.*

2. *(Sun, 2022b) Let $\mathscr{H}_m$ denote the set of unimprovable outcomes for sender m, then (i) $H \succeq_{MPS} H_{\mathscr{P}(\underline{\kappa},b)}$ for all $H \in \mathscr{H}_m$, and (ii) $H_{\mathscr{P}(\underline{\kappa},d)} \in \mathscr{H}_m$ for all $d \in [b, \overline{\kappa}]$.*

3. *(Sun, 2022b) Given any pure strategy profile of other senders π_{-m}, there exists some $d \in [b, r]$ such that the upper censorship policy $\mathscr{P}(\underline{\kappa}, d)$ is sender m's best response to π_{-m}.*

Observation F.3. *Let $U(\cdot)$ be sender m's utility function and suppose that $U(\cdot)$ is strictly inverse S-shaped on $[\underline{\kappa}, \overline{\kappa}]$ with inflection point ℓ. The following properties hold:*

1. *Kolotilin, Mylovanov and Zapechelnyuk (2022) The unique solution H to problem (MP') is induced by an lower censorship policy $\mathscr{P}(a, \overline{\kappa})$ with $a \leq \overline{\kappa}$.*

2. *(Sun, 2022b) Let $\mathscr{H}_m$ denote the set of unimprovable outcomes for sender m, then (i) $H \succeq_{MPS} H_{\mathscr{P}(a,\overline{\kappa})}$ for all $H \in \mathscr{H}_m$, and (ii) $H_{\mathscr{P}(c,\overline{\kappa})} \in \mathscr{H}_m$ for all $c \in [\underline{\kappa}, a]$.*

3. *(Sun, 2022b) Given any pure strategy profile of other senders π_{-m}, there exists some $c \in [\ell, a]$ such that the lower censorship policy $\mathscr{P}(c, \overline{\kappa})$ is sender m's best response to π_{-m}.*

The first statements of the two observations above are established by Kolotilin, Mylovanov and Zapechelnyuk (2022). These suggest that under monopolistic persuasion *upper* (resp. *lower*) censorship policies are uniquely optimal for a sender whose utility function $U(\cdot)$ is strictly S-shaped (resp. inverse S-shaped). The remaining statements, established by Sun (2022*b*), extend this insight to competition in persuasion. For a sender whose utility function is either strictly S-shaped or inverse S-shaped, any censorship policy that is no less informative than the monopolistic optimal one is unimprovable for him. Moreover, it is without loss of optimality for him to focus on a proper subset of censorship policies in the following sense: given any pure strategy profile of other senders, he can always find a best response from this subset of censorship policies. Notice that all information policies in this subset are no less informative than his monopolistically optimal one.

Finally, we introduce an easy-to-check sufficient condition for full disclosure to be the unique equilibrium outcome under competition in linear persuasion games. This condition is established by Sun (2022*a*).

Observation F.4. *([Sun, 2022a]) Let $\{U_m(\cdot)\}_{m\in M}$ be a profile of twice continuously differentiable utility functions defined on $[\underline{\kappa},\overline{\kappa}]$. Then full disclosure on $[\underline{\kappa},\overline{\kappa}]$ is the unique equilibrium outcome if there exists no interval $(x,y)\subset[\underline{\kappa},\overline{\kappa}]$ with $x<y$ such that $U''_m(k)\le 0$ for all $k\in(x,y)$ and $i\in M$.*

With these ingredients we are now ready to prove Theorems 3 and 4.

F.1 Proof of Theorem 3

For each $m\in M$ and $n\ge N_m$, recall from (BR) that

$$\mathscr{P}^m_n := \{\mathscr{P}(c,d):\ [a^m_n,b^m_n]\subseteq[c,d]\subseteq[\ell^m_n,r^m_n]\}$$

where a^m_n and b^m_n are the thresholds of the monopolistically optimal censorship policy. Clearly, $\mathscr{P}^m_n$ is a subset of censorship policies. We show below that for any pure strategy profile π_{-m} of other senders, there exists a $\pi_m\in\mathscr{P}^m_n$ such that π_m is sender m's best response to π_{-m}. We distinguish between three cases depending on the value of z^m_n.

If $z^m_n=-1$, then $\ell^m_n=a^m_n=0$ and $W^m_n(\cdot)$ is strictly S-shaped on $[-1,1]$ with inflection point $r^m_n\ge b^m_n$. By Observation F.2, for any π_{-m} there exists $d\in[b^m_n,r^m_n]$ such that $\pi_m=\mathscr{P}(-1,d)$ is sender m's best response to π_{-m}. Similarly, if $z^m_n=1$ then $r^m_n=b^m_n=1$ and $W^m_n(\cdot)$ is strictly inverse S-shaped on $[-1,1]$ with inflection point $\ell^m_n\le a^m_n$. Observation F.3 implies that for any π_{-m} there exists $c\in[\ell^m_n,a^m_n]$ such that $\pi_m=\mathscr{P}(c,1)$ is sender m's best response to π_{-m}. In both cases $\pi_m\in\mathscr{P}^m_n$ holds.

Finally, consider the case $z^m_n\in(-1,1)$ and let π_m be any best response to π_{-m} for sender m. Because the information environment is Blackwell-connected, the induced joint information policy $\langle\pi_m,\pi_{-m}\rangle$ must be unimprovable for sender m. Recall that $W^m_n(\cdot)$ satisfies the increasing slope property at point z^m_n (cf. Lemma 4), it thus follows from Lemma F.1 that $\langle\pi_m,\pi_{-m}\rangle$ must be Blackwell more informative than the cutoff policy $\mathscr{P}(z^m_n)$. This implies that there always exists a best response π_m that is Blackwell more informative than $\mathscr{P}(z^m_n)$ (i.e., $H_{\pi_m}\succeq_{MPS}H_{\mathscr{P}(z^m_n)}$).[17] For such π_m, it must be a best response to π_{-m} on both $[-1,z^m_n]$ and $[z^m_n,1]$ separately. By Lemma 5, $W^m_n(\cdot)$ is strictly inverse S-shaped on $[-1,z^m_n]$ with inflection point $\ell^m_n<z^m_n$ and strictly S-shaped on $[z^m_n,1]$ with inflection point $r^m_n>z^m_n$. It follows that there exists $c\in[\ell^m_n,a^m_n]$ and $d\in[b^m_n,r^m_n]$ such that $\mathscr{P}(c,z^m_n)$ is a best response to π_{-m} on $[-1,z^m_n]$ and $\mathscr{P}(z^m_n,d)$ is a best response on $[z^m_n,1]$. These together produce a censorship policy $\pi_m=\mathscr{P}(c,d)$, which belongs to $\mathscr{P}^m_n$, that is a best response to π_{-m}. This completes the proof.

[17] To see why, let $\pi=\langle\pi_m,\pi_{-m}\rangle$ and $\pi'=\langle\pi_m,\pi_{-m},\mathscr{P}(z^m_n)\rangle$. The best response property of π_m ensures that $H_\pi\succeq_{MPS}H_{\mathscr{P}(z^m_n)}$. Therefore, $H_\pi=H_{\pi'}$. Consider $\pi'_m=\langle\pi_m,\mathscr{P}(z^m_n)\rangle$ (which is always feasible) and observe that $H_{\pi'_m}\succeq_{MPS}H_{\mathscr{P}(z^m_n)}$ and $\pi'=\langle\pi'_m,\pi_{-m}\rangle$. Then π'_m must also be a best response to π_{-m} because $H_\pi=H_{\pi'}$.

F.2 Proof of Theorem 4

We first present two lemmas and use these to establish part (1) of the theorem. After that we turn to proving part (2), making use of a third lemma.

Lemma F.2. *Suppose Assumption 1 holds. Then any unimprovable outcome* $H \in \mathscr{H}$ *must satisfy* $H \succeq_{MPS} H_{\mathscr{P}(z_n^{\min}, z_n^{\max})}$*, where* $z_n^{\min} = \min_{m\in M}\{z_n^m\}$ *and* $z_n^{\max} = \max_{m\in M}\{z_n^m\}$.

Proof of Lemma F.2. By Lemma F.1, any $H \in \mathscr{H}$ must satisfy $H \succeq_{MPS} H_{\mathscr{P}(z_n^{\min})}$, $H \succeq_{MPS} H_{\mathscr{P}(z_n^{\max})}$, and be unimprovable on $[z_n^{\min}, z_n^{\max}]$. It then suffices to show that any unimprovable H must be fully revealing on $[z_n^{\min}, z_n^{\max}]$ to complete the proof. Recall from (C.7) that

$$W_n^{m\prime\prime}(k) = \left(2 - \phi_n^{m\prime}(k)\right)\hat{g}_n(k;q) + (k - \phi_n^m(k))\,\hat{g}_n'(k;q)$$

The first term is strictly positive for all $m \in M$ because part (2) of Assumption 1 ensures that $\phi_n^{m\prime}(k) < 2$ for all $k \in [z_n^{\min}, z_n^{\max}]$. Let I (resp. II) denote the index of the sender for whom $z_n^{\text{I}} = z_n^{\min}$ (resp. $z_n^{\text{II}} = z_n^{\max}$). Then for all $k \in [z_n^{\min}, z_n^{\max}]$ we have $\phi_n^{\text{I}}(k) \leq k \leq \phi_n^{\text{II}}(k)$. So, no matter what the sign of $\hat{g}_n'(k;q)$ is, $(k - \phi_n^m(k))\,\hat{g}_n'(k;q)$ must be non-negative for at least one $m \in \{\text{I}, \text{II}\}$. Hence, for any $k \in [z_n^{\min}, z_n^{\max}]$, $W_n^{m\prime\prime}(k) > 0$ must hold for at least one $m \in \{\text{I}, \text{II}\} \subseteq M$. Therefore, by Observation F.4, any unimprovable H must be fully revealing on $[z_n^{\min}, z_n^{\max}]$ and hence $H \succeq_{MPS} H_{\mathscr{P}(z_n^{\min}, z_n^{\max})}$. □

Combining Lemma F.1, Observations F.2 and F.3, and the curvature properties of $W_n^m(\cdot)$ summarized in Lemma 5, we obtain

Lemma F.3. *Suppose the single-crossing property holds for each sender* $m \in M$*. Then for any* $n \geq N_m$ *we have (i)* $H \succeq_{MPS} H_{\mathscr{P}(a_n^m, b_n^m)}$ *for all* $H \in \mathscr{H}_m$*, and (ii)* $\mathscr{P}(a,b) \in \mathscr{H}_m$ *for all* $a \in [-1, a_n^m]$ *and* $b \in [b_n^m, 1]$.

In words, Lemma F.3 says that under the single-crossing property and for sufficiently large n all unimprovable outcomes for sender m must be no less informative than his monopolistically optimal censorship policy. Moreover, all censorship policies that are more informative than the monopolistically optimal one are unimprovable for sender m.

We now use Lemmas F.2 and F.3 to establish part (1) of Theorem 4. Consider any $n \geq N$. By Lemma F.2, $H \succeq_{MPS} H_{\mathscr{P}(z_n^{\min}, z_n^{\max})}$ must hold for all $H \in \mathscr{H} = \cap_{m\in M}\mathscr{H}_m$. By Lemma F.3, $H \succeq_{MPS} H_{\mathscr{P}(a_n^m, b_n^m)}$ must hold for all $H \in \mathscr{H}_m$. Moreover, for each $m \in M$, it holds that $H_{\mathscr{P}(c,d)} \in \mathscr{H}_m$ for all $c \in [-1, a_n^m]$ and $d \in [b_n^m, 1]$. Therefore, $\mathscr{P}(a_n^{\min}, b_n^{\max})$ is unimprovable for all senders and hence $H_{\mathscr{P}(a_n^{\min}, b_n^{\max})} \in \mathscr{H}$. Next we show that any $H \in \mathscr{H}$ must be weakly more informative than $\mathscr{P}(a_n^{\min}, b_n^{\max})$, that is $H \succeq_{MPS} H_{\mathscr{P}(a_n^{\min}, b_n^{\max})}$. Let $\tilde{i}$ (resp. $\tilde{j}$) denote the identity of the sender with $a_n^{\tilde{i}} = a_n^{\min}$ (resp. $b_n^{\tilde{j}} = b_n^{\max}$). Recall that any $H \in \mathscr{H}$ must satisfy $H \succeq_{MPS} H_{\mathscr{P}(a_n^m, b_n^m)}$ with

$a_n^m \leq z_n^m \leq b_n^m$ for all $m \in M$. The choices of $\tilde{i}$ and $\tilde{j}$ imply that $\left[a_n^{\tilde{i}}, b_n^{\tilde{i}}\right]$, $[z_n^{\min}, z_n^{\max}]$ and $\left[a_n^{\tilde{j}}, b_n^{\tilde{j}}\right]$ are overlapping and $\left[a_n^{\tilde{i}}, b_n^{\tilde{i}}\right] \cup [z_n^{\min}, z_n^{\max}] \cup \left[a_n^{\tilde{j}}, b_n^{\tilde{j}}\right] = [a_n^{\min}, b_n^{\max}]$. Therefore, $H \succeq_{MPS} H_{\mathscr{P}(a_n^{\min}, b_n^{\max})}$ must hold for all $H \in \mathscr{H}$ and this completes the proof for statement (1) of Theorem 4.

Next we prove statement (2) of Theorem 4. Given any censorship policy $\mathscr{P}(c,d)$ with revelation interval $[c,d] \subseteq [-1,1]$, each sender m's expected payoff under this policy is

$$\mathbb{E}_{\mathscr{P}(c,d)}\left[W_n^m(\cdot)\right] = F(c) W_n^m\left(\underline{\mu}_F(c)\right) + \int_c^d W_n^m(k) dF(k) \left(1 - F(d)\right) W_n^m\left(\overline{\mu}_F(d)\right) \tag{F.1}$$

where $\underline{\mu}_F(c) := \mathbb{E}_F[k | k < c]$ and $\overline{\mu}_F(d) := \mathbb{E}_F[k | k > d]$. Lemma F.4 shows that $\mathbb{E}_{\mathscr{P}(c,d)}[W_n^m(\cdot)]$ is single-peaked in both thresholds c and d.

Lemma F.4. *The following properties hold:*
(i) $\frac{\partial \mathbb{E}_{\mathscr{P}(c,d)}[W_n^m(\cdot)]}{\partial d} > (<) 0$ *for* $d < (>) b_n^m$*, and*
(ii) $\frac{\partial \mathbb{E}_{\mathscr{P}(c,d)}[W_n^m(\cdot)]}{\partial c} > (<) 0$ *for* $c > (<) a_n^m$.

Proof of Lemma F.4. Taking derivatives of (F.1) with respect to c and d yield[18]

$$\begin{aligned}
\frac{\partial \mathbb{E}_{\mathscr{P}(c,d)}\left[W_n^m(\cdot)\right]}{\partial d} &= f(d)\left(W_n^m(d) - W_n^m\left(\overline{\mu}_F(d)\right)\right) + \left(1 - F(d)\right) W_n^{m\prime}\left(\overline{\mu}_F(d)\right) \overline{\mu}_F'(d) \\
&= f(d) \cdot \left(\overline{\mu}_F(d) - d\right) \cdot \left[\frac{W_n^m\left(\overline{\mu}_F(d)\right) - W_n^m(d)}{\overline{\mu}_F(d) - d} - W_n^{m\prime}\left(\overline{\mu}_F(d)\right)\right] \\
\frac{\partial \mathbb{E}_{\mathscr{P}(c,d)}\left[W_n^m(\cdot)\right]}{\partial c} &= f(c)\left(W_n^m\left(\underline{\mu}_F(c)\right) - W_n^m(c)\right) + F(c) \cdot W_n^{m\prime}\left(\underline{\mu}_F(c)\right) \underline{\mu}_F'(c) \\
&= f(c) \cdot \left(c - \underline{\mu}_F(c)\right) \cdot \left[W_n^{m\prime}\left(\underline{\mu}_F(c)\right) - \frac{W_n^m(c) - W_n^m\left(\underline{\mu}_F(c)\right)}{c - \underline{\mu}_F(c)}\right]
\end{aligned}$$

Because both $\tilde{f}(d)$ and $\overline{\mu}_F(d) - d$ are positive, $\frac{\partial \mathbb{E}_{\mathscr{P}(c,d)}[W_n^m(\cdot)]}{\partial d}$ is sign-equivalent to

$$\frac{W_n^m\left(\overline{\mu}_F(d)\right) - W_n^m(d)}{\overline{\mu}_F(d) - d} - W_n^{m\prime}\left(\overline{\mu}_F(d)\right) \tag{F.2}$$

By Lemma 5, $W_n^m(\cdot)$ is strictly S-shaped on $[z_n^m, 1]$ and hence $\frac{W_n^m(\overline{\mu}_F(d)) - W_n^m(d)}{\overline{\mu}_F(d) - d} - W_n^{m\prime}\left(\overline{\mu}_F(d)\right)$ crosses zero at most once and above at b_n^m, which is pinned down by condition (FOC: b_n). This proves part (i). The proof for part (ii) is similar; it exploits the inverse S-shape property of $W_n^m(\cdot)$ on $[-1, z_n^m]$ and the definition of a_n^m. □

We establish below that any equilibrium outcome in pure and weakly undominated strategies

[18] In the derivation we exploit the fact that $\overline{\mu}_F'(d) = \frac{f(d)}{1-F(d)}\left(\overline{\mu}_F(d) - d\right)$ and $\underline{\mu}_F'(c) = \frac{f(c)}{F(c)}\left(c - \underline{\mu}_F(c)\right)$.

must be both no more and no less informative than $\mathscr{P}(a_n^{\min}, b_n^{\max})$. These together imply the uniqueness of $\mathscr{P}(a_n^{\min}, b_n^{\max})$ as the induced outcome of any pure strategy equilibrium in weakly undominated strategies.

We first show for any $m \in M$ that all $\mathscr{P}(c,d)$ with $d > b_n^{\max}$ are weakly dominated by $\mathscr{P}(c, b_n^{\max})$, provided that all other senders $i \neq m$ choose strategies from $\mathscr{P}_n^i$. Let π_{-m} denote a strategy profile by other senders and under π_{-m} let $\eta \in [b_n^{\max}, 1]$ be the threshold such that $k \in [b_n^{\max}, \eta]$ are revealed, while $k > \eta$ are censored. Replacing $\mathscr{P}(c,d)$ with $\mathscr{P}(c, b_n^{\max})$ can only make a difference in states $k \in [b_n^{\max}, 1]$. If $d \leq \eta$, then such replacement has no effect on the joint information policy so sender m is indifferent to it. If instead $d > \eta$, then such replacement lowers the threshold of the upper censoring interval and thus reduces the informativeness of the joint policy. By Lemma F.4, for each $m \in M$, $\mathbb{E}_{\mathscr{P}(c,d)}[W_n^m(\cdot)]$ is single-peaked in d with a peak at b_n^m. Since $\eta \geq b_n^{\max} > b_n^m$, it follows that any sender m's expected payoff would increase were $\mathscr{P}(c,d)$ replaced by $\mathscr{P}(c, b_n^{\max})$. Hence, any $\mathscr{P}(c,d)$ with $d > b_n^{\max}$ is weakly dominated by $\mathscr{P}(c, b_n^{\max})$. Using an analogous argument we can also show that any $\mathscr{P}(c,d)$ with $c < a_n^{\min}$ is weakly dominated by $\mathscr{P}(a_n^{\min}, d)$. Together these imply that any $\mathscr{P}(c,d)$ with $d > b_n^{\max}$ or $c < a_n^{\min}$ is weakly dominated. This shows that any outcome induced by a pure-strategy equilibrium with undominated strategies must be weakly less informative than $\mathscr{P}(a_n^{\min}, b_n^{\max})$.

Next we show that no equilibrium outcome can be strictly less informative than censorship policy $\mathscr{P}(a_n^{\min}, b_n^{\max})$.[19] Observe that the structure of $\mathscr{P}_n^m$ implies that any feasible outcome must be weakly more informative than $\mathscr{P}(a_n^m, b_n^m)$ for all $m \in M$. Therefore, if $\cup_{m \in M}[a_n^m, b_n^m] = \left[a_n^{\min}, b_n^{\max}\right]$ the result holds trivially. In what follows we assume that $\cup_{m \in M}[a_n^m, b_n^m]$ is a proper subset of $\left[a_n^{\min}, b_n^{\max}\right]$. In this case, there must be at least one pair of senders $l, r \in M$ such that (i) $b_n^l < a_n^r$, and (ii) for all $m \in M \setminus \{l, r\}$ there are $[a_n^m, b_n^m] \cap \left(b_n^l, a_n^r\right) = \emptyset$.[20] By the construction of $\mathscr{P}_n^m$ for all $m \in M$, there could be at most one nontrivial pooling interval $(x,y) \subseteq \left(b_n^l, a_n^r\right)$. Fix this pooling interval (x,y) and let $\mu(x,y) := \mathbb{E}_F[k | k \in (x,y)]$, we obtain that the expected utility of sender $m = \{l, r\}$ conditional on event $k \in \left(b_n^l, a_n^r\right)$ is

$$V_m = \int_{b_n^l}^{x} W_n^m(k) d\widetilde{F}(k) + \left(\widetilde{F}(y) - \widetilde{F}(x)\right) W_n^m(\mu(x,y)) + \int_{y}^{a_n^r} W_n^m(k) d\widetilde{F}(k)$$

where $\widetilde{F}(\cdot)$ denotes the cdf of the distribution of k conditional on $k \in \left[b_n^l, a_n^r\right]$. Taking derivatives of

[19] Note that the information environment is no longer Blackwell-connected if each sender m is restricted to choose information policies from $\mathscr{P}_n^m$ only. Therefore, Proposition 2 of Gentzkow and Kamenica (2017*b*) no longer applies (i.e., a feasible outcome being unimprovable to all senders is no longer necessary for that outcome to be an equilibrium).

[20] In fact, there could be at most $|M| - 1$ such pairs. The argument presented below holds for any such pair.

V_m with respect to x and y yields[21]

$$\begin{aligned}\frac{\partial V_m}{\partial x} &= \widetilde{f}(x)\left(W_n^m(x) - W_n^m(\mu(x,y))\right) + \left(\widetilde{F}(y) - \widetilde{F}(x)\right) W_n^{m\prime}(\mu(x,y))\,\mu_x(x,y) \\ &= \widetilde{f}(x)\left(\mu(x,y) - x\right)\left[W_n^{m\prime}(\mu(x,y)) - \frac{W_n^m(\mu(x,y)) - W_n^m(x)}{\mu(x,y) - x}\right]\end{aligned}$$

and

$$\begin{aligned}\frac{\partial V_m}{\partial y} &= \widetilde{f}(y)\left(W_n^m(\mu(x,y)) - W_n^m(y)\right) + \left(\widetilde{F}(y) - \widetilde{F}(x)\right) W_n^{m\prime}(\mu(x,y))\,\mu_y(x,y) \\ &= \widetilde{f}(x)\left(y - \mu(x,y)\right)\left[W_n^{m\prime}(\mu(x,y)) - \frac{W_n^m(y) - W_n^m(\mu(x,y))}{y - \mu(x,y)}\right]\end{aligned}$$

For both l and r to have no incentive to reveal any extra information, it is necessary that $\frac{\partial V_l}{\partial x} \leq 0$ and $\frac{\partial V_r}{\partial y} \geq 0$, or equivalently[22]

$$W_n^{l\prime}(\mu(x,y)) \leq \frac{W_n^l(\mu(x,y)) - W_n^l(x)}{\mu(x,y) - x} \tag{F.3}$$

and

$$W_n^{r\prime}(\mu(x,y)) \geq \frac{W_n^r(y) - W_n^l(\mu(x,y))}{y - \mu(x,y)} \tag{F.4}$$

Because $z_n^l \leq b_n^l \leq x < y \leq a_n^r \leq z_n^r$, both $[z_n^l, 1]$ and $[-1, z_n^r]$ must contain (x,y) in their interior. For (F.3) to hold, $\mu(x,y) > r_n^l$ must be true so that $\mu(x,y)$ falls into the concave region of $W_n^l(\cdot)$. Similarly, for (F.4) to hold, $\mu(x,y) < \ell_n^r$ must hold for $\mu(x,y)$ to fall into the concave region of $W_n^r(\cdot)$. These together imply that both $W_n^l(\cdot)$ and $W_n^r(\cdot)$ are strictly concave on an open neighborhood of $\mu(x,y)$, which lies in $[z_n^l, z_n^r]$. This is, however, impossible because the proof of Lemma F.2 shows that for any $k \in [z_n^l, z_n^r]$, $W_n^{m\prime\prime}(k) > 0$ must hold for at least one $m \in \{l, r\}$.[23] Therefore, the incentive compatibility conditions for senders l and r cannot be simultaneously satisfied and hence it is impossible to have any non-trivial pooling interval (x,y) in equilibrium. This implies that no equilibrium outcome can be strictly less informative than $\mathscr{P}(a_n^{\min}, b_n^{\max})$ and thus completes the proof for statement (2) of Theorem 4.

[21] Here we use the fact that $\mu_x(x,y) := \frac{\partial \mu}{\partial x} = \frac{\widetilde{f}(x)(\mu(x,y)-x)}{\widetilde{F}(y)-\widetilde{F}(x)}$ and $\mu_y(x,y) := \frac{\partial \mu}{\partial y} = \frac{\widetilde{f}(y)(y-\mu(x,y))}{\widetilde{F}(y)-\widetilde{F}(x)}$.

[22] This is because each $m \in \{l, r\}$ can only choose censorship policies from $\mathscr{P}_n^m$, whose revelation interval must contain $[a_n^m, b_n^m]$. Therefore, since $(x,y) \subseteq \left(b_n^l, a_n^r\right)$, only sender l can marginally increase x while only sender r can marginally decrease y. These two inequalities ensure that such marginal deviations are not profitable for either sender.

[23] This is an application of Lemma F.2 with $M = \{l, r\}$.

References for Appendices

Bagnoli, Mark and Ted Bergstrom. 2005. "Log-concave probability and its applications." *Economic theory* 26(2):445–469.

Dworczak, Piotr and Giorgio Martini. 2019. "The simple economics of optimal persuasion." *Journal of Political Economy* 127(5):1993–2048.

Gentzkow, Matthew and Emir Kamenica. 2017. "Competition in Persuasion." *The Review of Economic Studies* 84(1):300–322.

Groeneveld, Richard and Glen Meeden. 1977. "The mode, median, and mean inequality." *The American Statistician* 31(3):120–121.

Kolotilin, Anton, Timofiy Mylovanov and Andriy Zapechelnyuk. 2022. "Censorship as optimal persuasion." *Theoretical Economics* 17(2):561–585.

Krishna, Vijay. 2009. *Auction Theory*. Academic press.

Sun, Junze. 2022*a*. "Full information disclosure in linear persuasion games." Working paper.

Sun, Junze. 2022*b*. "On the optimality of censorship in linear persuasion games." Working paper.

Van Zwet, Willem R. 1980. "A strong law for linear functions of order statistics." *The Annals of Probability* 8(5):986–990.