CENSORSHIP AS OPTIMAL PERSUASION

Anton Kolotilin, Tymofiy Mylovanov, Andriy Zapechelnyuk

Abstract. A sender designs a signal about the state of the world to persuade a receiver. Under standard assumptions, an optimal signal censors states on one side of a cutoff and reveals all other states. This result holds in continuous and discrete environments with general and monotone partitional signals. The sender optimally censors more information if she is more biased, if she is more certain about the receiver's preferences, and if the receiver is easier to persuade. We apply our results to the problem of media censorship by a government.

JEL Classification: D82, D83, L82

Keywords: Bayesian persuasion, information design, censorship, media

Date: December 10, 2019.

Kolotilin: School of Economics, UNSW Australia, Sydney, NSW 2052, Australia. *E-mail:* akolotilin@gmail.com.
Mylovanov: Cabinet of Ministers of Ukraine, Government of Ukraine; University of Pittsburgh, Department of Economics, 4714 Posvar Hall, 230 South Bouquet Street, Pittsburgh, PA 15260, USA. *E-mail:* mylovanov@gmail.com.
Zapechelnyuk: School of Economics and Finance, University of St Andrews, Castlecliffe, the Scores, St Andrews KY16 9AR, UK. *E-mail:* az48@st-andrews.ac.uk.

We are grateful for discussions with Ming Li with whom we worked on the companion paper Kolotilin et al. (2017). An early version of the results in this paper and the results in the companion paper were presented in our joint working paper Kolotilin et al. (2015). We thank Ricardo Alonso, Dirk Bergemann, Patrick Bolton, Alessandro Bonatti, Steven Callander, Odilon Câmara, Rahul Deb, Péter Esö, Johannes Hörner, Florian Hoffman, Roman Inderst, Emir Kamenica, Navin Kartik, Daniel Kräehmer, Hongyi Li, Marco Ottaviani, Mallesh Pai, Andrea Prat, Ilya Segal, Larry Samuelson, Joel Sobel, Konstantin Sonin, as well as many conference and seminar participants, for helpful comments. Kolotilin gratefully acknowledges financial support from the Australian Research Council Discovery Early Career Research Award DE160100964 and from MIT Sloan's Program on Innovation in Markets and Organizations. Zapechelnyuk gratefully acknowledges financial support from the Economic and Social Research Council Grant ES/N01829X/1.

1. Introduction

There has been a rapid growth of the literature on Bayesian persuasion and information design over the past decade. The range of applications in this literature is impressive and includes clinical trials, bank stress tests, school grading policies, quality certification, advertising strategies, transparency in organizations, persuasion of voters, and media control. In many of these applications, it is optimal to disclose information about an uncertain state of the world via a *censorship signal* that censors states on one side of a cutoff and reveals states on the other side of the cutoff. Fully informative and completely uninformative signals are special cases of censorship.

We characterize necessary and sufficient conditions under which a censorship signal is optimal in the *linear persuasion* problem, where the sender and receiver's utilities depend on the beliefs about the state only through the expected state. We thus unify and generalize sufficient conditions scattered across the literature. In addition, we establish new results for environments where the state is discrete and a signal is constrained to be *monotone partitional*. These environments are relevant in applications and take into account constraints faced by information designers in practice.[1] Finally, we provide monotone comparative statics results on the informativeness of the optimal signal.

The linear persuasion problem is a tractable workhorse model in the persuasion literature. This problem can be described by the prior distribution of the state and the sender's *indirect utility* function of the expected state. We consider a slightly different but equivalent formulation of the problem, which is easier to interpret.

In our model, the receiver chooses one of two actions: to accept or reject a proposal. If the receiver rejects the proposal, the sender and receiver's utilities are normalized to zero. If the receiver accepts the proposal, the sender and receiver's utilities depend on the unknown state of the world and the receiver's private type. The state and type are independent random variables that represent, respectively, the receiver's benefit and cost from acceptance. We assume that the type is a continuous random variable, but we allow the state to be either a continuous or discrete random variable.[2]

The sender's utility is linear in the state and is not perfectly aligned with the receiver's utility. An important *canonical case* of the model is where the sender's utility is a weighted sum of the receiver's utility and action.

To influence the receiver's action, the sender designs a signal that reveals information about the state. After observing the signal realization, the receiver updates his beliefs about the state and chooses an action that maximizes his expected utility.

[1]For example, a credit rating of financial institutions is typically a monotone partition. Neither nonmonotone partitions that pool high- and low-performing institutions nor stochastic disclosure rules that map an institution's financial performance to a random rating appear to be feasible.

[2]Extending the analysis to the case where the state and type are general random variables presents technical difficulties but does not yield new economic insights.

The main result shows that if the sender's indirect utility is S-shaped in the expected state (that is, convex below some threshold and concave above the threshold), then, and only then, an *upper-censorship* signal that pools states above a cutoff and reveals states below the cutoff is optimal for all prior distributions of the state.[3] In the canonical case, the main result shows that if the probability density of receiver types is log-concave, then, and only then, upper-censorship is optimal for all prior distributions of the state and for all weights that the sender puts on the receiver's utility and action.[4]

The main result is obtained under two scenarios. In the first scenario, the signal is a random variable with arbitrary correlation with the state. In the second scenario, the signal is constrained to be a monotone partition of the state space, so that each state is either revealed or pooled with some adjacent states.

To compare the main result under these two scenarios, consider the cutoff of an optimal upper-censorship signal. Notice that upper-censorship does not specify whether the cutoff state is pooled or revealed. If the state is a continuous random variable, then it does not matter whether this cutoff state is pooled or revealed. Thus, an optimal upper-censorship signal is, in fact, a monotone partition. However, if the state is a discrete random variable, then the optimal signal is *stochastic* upper censorship so that the cutoff state is generally pooled and revealed with interior probabilities. In contrast, the optimal monotone partition is *deterministic* upper censorship so that the cutoff state is either completely pooled or fully revealed.

In many applications of interest, the state is discrete, and thus optimal signals and optimal monotone partitions generally differ. An optimal upper-censorship signal provides an upper bound on the value of persuasion. To achieve this bound, the sender should be able to commit to randomize over two signal realizations conditional on the cutoff state, which may be hard to enforce in practice. In contrast, monotone partitions are relatively simple signals that can enforced and verified ex post.

The problem of finding an optimal monotone partition is a discrete optimization problem, which cannot be solved using existing tools from the persuasion literature. Nevertheless, we show that if the sender's indirect utility is S-shaped (or, in the canonical case, the density of receiver type is log-concave), then upper censorship is an optimal monotone partition. To prove this result, for any monotone partition different from upper censorship, we explicitly construct a deterministic upper-censorship signal that is preferred by the sender.

When upper censorship is optimal, it is possible to perform the comparative statics analysis on the informativeness of the optimal signal. We show that, in the canonical case, the sender optimally censors more information (i) if the sender is more biased

[3]Analogously, a *lower-censorship* signal that pools states below a cutoff is optimal for all prior distributions if and only if the sender's indirect utility has an inverted S-shape.

[4]Most commonly used probability densities are log-concave. Log-concave densities exhibit nice properties, such as single-peakedness and the monotonicity of the hazard rate.

in that the sender puts a smaller weight on the receiver's utility, (ii) if the receiver is easier to persuade in that the receiver's cost of accepting the proposal is smaller, and (iii) if the sender is more certain about the receiver's preferences in that the density of receiver types is more concentrated around its mode.

We apply our results to the problem of media censorship by the government. We consider a stylized setting with a finite number of media outlets and a continuum of heterogeneous citizens (receivers). Each media outlet is identified by its editorial policy that specifies an interval of *endorsement* states. We permit the aggregate action of the citizens to affect their utility but not their optimal actions. For example, an election outcome impacts all citizens but does not change their preferences over candidates. The government wishes to influence citizens' actions by deciding which media outlets to censor. In the canonical case, if the probability density of citizens types is log-concave, then the optimal media censorship policy is to permit all sufficiently loyal media outlets and to censor the remaining outlets. Our comparative statics results suggest that the government should optimally increase censorship (i) if influencing society decisions becomes relatively more important than maximizing individual welfare, (ii) if the society experiences an ideology shock in favor of the government, and (iii) if the society becomes less diverse in ideology and taste.

Related Literature. The literature on Bayesian persuasion was set in motion by the seminal papers of Rayo and Segal (2010) and Kamenica and Gentzkow (2011). The linear case, which is the subject of this paper, is prevalent in the literature.[5] Kolotilin (2018) provides necessary and sufficient *prior dependent* conditions for the optimality of censorship in the case of a continuous state.[6] Alonso and Câmara (2016b) provide sufficient conditions in the case of a discrete state. In contrast, our paper provides necessary and sufficient conditions that are *prior independent* (as well as *bias independent* in the canonical case) and apply in both continuous and discrete cases. More importantly, we show that these conditions are also necessary and sufficient for *deterministic* censorship to be optimal in the class of *monotone partitional* signals (simple signals). There are no counterparts to this result in the literature.

Similar to us, Alonso and Câmara (2016b) show that the optimal signal is less informative if the receiver is easier to persuade. Their proof is simpler, but it is valid only when the sender's utility is state-independent. But our comparative statics results with respect to the sender's bias and uncertainty are novel.

Censorship policies commonly emerge as optimal signals in linear persuasion models in a variety of contexts, starting from the prosecutor-judge example, as well as lobbying and product advertising examples, in Kamenica and Gentzkow (2011). Other contexts

[5]Notable exceptions include Rayo and Segal (2010), Goldstein and Leitner (2018), Guo and Shmaya (2019), and Kolotilin and Wolitzky (2019).

[6]See also Gentzkow and Kamenica (2016), Kolotilin, Mylovanov, Zapechelnyuk, and Li (2017), and Dworczak and Martini (2019) for sufficient conditions.

where censorship is optimal include grading policies (Ostrovsky and Schwarz, 2010), media control (Gehlbach and Sonin, 2014; Ginzburg, 2019), clinical trials (Kolotilin, 2015), voter persuasion (Alonso and Câmara, 2016a,b), transparency benchmarks (Duffie, Dworczak, and Zhu, 2017), stress tests (Goldstein and Leitner, 2018; Orlov, Zryumov, and Skrzypach, 2018), trading mechanisms (Romanyuk and Smolin, 2019; Dworczak, 2017; Yamashita, 2018), quality certification (Zapechelnyuk, 2019), and relational communication (Kolotilin and Li, 2019).

Our application to media censorship fits into the literature on media capture that addresses the problem of media control by governments, political parties, or lobbying groups. Besley and Prat (2006) pioneered this literature by studying how a government's incentives to censor free media depend on plurality (the number of media outlets) and on transaction costs of bribing the media. In their model, media outlets are identical and all voters have common interests. Thus, the government's optimal decision reduces to either capturing all media outlets or none. Relatedly, Gehlbach and Sonin (2014) consider a setting with a government-influenced monopoly media outlet that exploits the trade-off between the government's objective to "mobilize" the population for some collective goal and to collect the revenue from subscribers who demand informative news. Gehlbach and Sonin (2014) show that the presence of the "mobilization" objective increases the news bias, whereas the subscription revenues reduce the bias but can cause the government to nationalize the media outlet.[7]

The models of Besley and Prat (2006) and Gehlbach and Sonin (2014) have two states of the world and either a single media outlet or a few identical media outlets. Hence, the government uses the same censorship policy for each outlet. We are the first to consider a richer model of media censorship with a continuum of states and multiple heterogeneous media outlets. As a result, the government optimally discriminates media outlets, by permitting sufficiently loyal ones and banning the remaining ones.

2. Model

2.1. **Setup.** There are two players: a sender (she) and a receiver (he). The receiver chooses whether to accept a proposal ($a = 1$) or reject it ($a = 0$). The proposal has an uncertain value $\omega \in [0,1]$. By accepting the proposal, the receiver forgoes an outside option worth $r \in [0,1]$, so the receiver's utility is $a(\omega - r)$. The sender's utility is $av(\omega, r)$, where $v(\omega, r)$ is linear in ω and continuously differentiable in r. We will refer to ω as *state* and to r as *type*, and denote their distributions by F and G, respectively. Throughout the paper we assume that distribution G of r has a continuously differentiable and strictly positive density g.[8]

[7] In different contexts, media control by a government has also been studied by Egorov, Guriev, and Sonin (2009), Edmond (2013), and Lorentzen (2014). See also the overview of the literature on media capture, slant, and transparency in Prat and Strömberg (2013).

[8] This assumption is made for clarity of exposition. The results can be extended to general G.

As in many applications the state is either a continuous or discrete random variable, we will separately analyze these two cases. We say that the state ω and distribution F are *continuous* if F has a density f on $[0,1]$. We say that the state ω and distribution F are *discrete* if the support of F is a finite set.

A particularly tractable and intuitive special case is where the sender's utility is a weighted sum of the receiver's utility and action. We call this the *canonical case*. Formally, in the canonical case, the sender's utility when the proposal is accepted is

$$v(\omega, r) = 1 + \rho(\omega - r), \quad \rho \in \mathbb{R}. \tag{A$_1$}$$

That is, the sender is biased towards $a = 1$ but also puts a weight ρ on the receiver's utility. In particular, if the weight ρ is large, then the sender and receiver's interests are aligned, whereas if the weight is zero, then the sender cares only about the receiver's action.

The receiver privately knows his type, but he does not observe the state. The sender can influence the action taken by the receiver, $a = 1$ or $a = 0$, by releasing a signal that reveals information about the state. A *signal* is a random variable $s \in [0,1]$ that is independent of r but is, possibly, correlated with ω. For example, s is fully informative if it is perfectly correlated with ω, and s is completely uninformative if it is independent of ω.

The timing is as follows. First, the sender publicly chooses a signal s. Then, realizations of ω, r, and s are drawn. Finally, the receiver observes the realizations of his type r and the signal s, and then chooses between $a = 0$ and $a = 1$.

We consider two scenarios:

A. The sender can choose any signal.

B. The sender can choose any monotone partitional signal.

Formally, a signal s is a *monotone partitional signal* if there exists a nondecreasing function $\xi(\omega)$ such that $s = \xi(\omega)$ for each ω. Every such signal induces a partition of the state space $[0,1]$ into intervals and singletons, and the receiver observes the partition element that contains the state.

Under Scenario A, we are interested in an *optimal signal* that maximizes the sender's expected utility among all signals. This is a standard persuasion problem.

Under Scenario B, we are interested in an *optimal monotone partition* that maximizes the sender's expected utility among all monotone partitional signals. This scenario incorporates constraints that information designers often face in practice. For example, a non-monotone grading policy that gives better grades to worse performing students will be perceived as unfair and will be manipulated by strategic students.

2.2. **Upper Censorship.** A subset of signals called upper censorship will play a special role in this paper.

An *upper-censorship signal* reveals states below a specified cutoff and pools states above this cutoff. When the state is equal to the cutoff, it can be revealed or pooled. If the state is continuous, it does not matter whether the cutoff state is revealed or pooled, because this is a zero probability event. However, if the state is discrete, we will distinguish between deterministic and stochastic upper-censorship signals, depending on what happens at the cutoff state.

A *lower-censorship signal* is defined symmetrically: it pools states below a specified cutoff and reveals states above this cutoff. Because the case of lower censorship is symmetric (see Remark 1 below), for clarity of exposition, we focus on upper censorship throughout the paper.

A signal s is *stochastic upper censorship* if there exists a cutoff pair (ω^*, q^*) consisting of a cutoff state $\omega^* \in [0,1]$ and probability $q^* \in [0,1]$ such that states below ω^* are revealed, states above ω^* are pooled, and the state $\omega = \omega^*$ is revealed with probability q^* and pooled with probability $1-q^*$. For example, s can be expressed as

$$s = \begin{cases} \omega \text{ with probability one}, & \text{if } \omega < \omega^*, \\ \omega^* \text{ and } m^* \text{ with probabilities } q^* \text{ and } 1-q^*, & \text{if } \omega = \omega^*, \\ m^* \text{ with probability one}, & \text{if } \omega > \omega^*, \end{cases}$$

where

$$m^* = \frac{\int_{(\omega^*,1]} \omega \mathrm{d}F(\omega) + \omega^*(1-q^*)\Pr[\omega = \omega^*]}{\int_{(\omega^*,1]} \mathrm{d}F(\omega) + (1-q^*)\Pr[\omega = \omega^*]} \tag{1}$$

is the posterior expected state induced by the pooling signal. Notice that m^* is a function of ω^* and q^*.

A stochastic upper-censorship signal s with a cutoff pair (ω^*, q^*) is *deterministic upper censorship* if and only if $q^* \in \{0,1\}$, in which case s can be expressed as a monotone partition of the state space $[0,1]$. For example, if $q^* = 0$, then the signal s fully reveals states $\omega < \omega^*$ and pools states $\omega \geq \omega^*$, so s can be expressed as

$$s = \begin{cases} \omega, & \text{if } \omega < \omega^*, \\ m^*, & \text{if } \omega \geq \omega^*, \end{cases}$$

where

$$m^* = \mathbb{E}[\omega | \omega \geq \omega^*]$$

is the expected state conditional on being at least ω^*. Note that the fully informative signal and the completely uninformative signal are deterministic upper-censorship signals with $(\omega^*, q^*) = (1,1)$ and $(\omega^*, q^*) = (0,0)$, respectively.

2.3. **Benchmark.** To illustrate the difference between the cases of a continuous and discrete state, as well as between Scenarios A and B, we solve a benchmark example where there is no uncertainty about the receiver's type.

Let the state ω have the expected value $\mathbb{E}[\omega] = \frac{1}{2}$, and let the receiver's type r be known to the sender and satisfy $r \in \left(\frac{1}{2}, 1\right)$. In addition, suppose that the sender wishes to maximize the probability that the receiver accepts the proposal, so $v(\omega, r) = 1$ for all ω and all r.

Observe that if the sender reveals no information about ω, the receiver will evaluate ω by its expected value of $\frac{1}{2}$, thus having utility $\frac{1}{2} - r < 0$ from accepting the proposal. So, revealing no information makes the receiver reject the proposal with certainty. The sender can do better by fully revealing ω, in which case the proposal is accepted with probability $\Pr[\omega \geq r]$.

However, the sender can do even better by pooling states above r with states below r while keeping the expected state induced by the pooling signal at least r. Thus, the receiver who is unable to distinguish between the states in the pool will accept the proposal at all of these states.

To illustrate the case of the continuous state, suppose that ω is uniformly distributed on $[0, 1]$. The largest pool that maintains the posterior expectation at least r is the interval $[2r - 1, 1]$. Indeed, once the receiver learns that $\omega \in [2r - 1, 1]$, given the uniform prior of ω, the posterior expected state is r. So the proposal is accepted when $\omega \geq 2r - 1$ and rejected when $\omega < 2r - 1$.[9] Thus, the deterministic upper-censorship signal with cutoff $\omega^* = 2r - 1$ is optimal. The sender's expected utility is equal to

$$\Pr[\text{proposal is accepted}] = \Pr[\omega \geq 2r - 1] = 2(1 - r),$$

which is twice as high as that from the fully informative signal:

$$\Pr[\text{proposal is accepted}] = \Pr[\omega \geq r] = 1 - r.$$

When ω is a discrete random variable, an optimal way to reveal information about ω takes the form of stochastic upper censorship. For illustration, suppose that ω can only be 0 or 1, equally likely. Now, simply pooling states is not helpful for the sender. Pooling 0 and 1 yields the posterior expected state of $\frac{1}{2}$, which is smaller than the type r, resulting in a rejection of the proposal. Including any other states in the pool makes no difference as these states never occur. Thus, an optimal monotone partition is the fully informative signal and the probability that the receiver accepts the proposal is simply equal to the probability that $\omega = 1$:

$$\Pr[\text{proposal is accepted}] = \Pr[\omega \geq r] = \Pr[\omega = 1] = \frac{1}{2}.$$

The sender can do strictly better by partial (stochastic) pooling. If $\omega = 1$, let the signal realization be "accept" with certainty. If $\omega = 0$, let the signal realization be "reject" with some probability q and "accept" with probability $1 - q$. So, when the receiver observes "accept", the posterior expected state is

$$\mathbb{E}[\omega | \text{“accept”}] = \frac{\Pr[\omega = 1]}{\Pr[\omega = 1] + \Pr[\omega = 0] \cdot (1 - q)} = \frac{1/2}{1/2 + 1/2 \cdot (1 - q)} = \frac{1}{2 - q}.$$

[9] Whether states strictly below ω^* are revealed or pooled among themselves does not matter, because they are all strictly below r and, thus, induce the receiver to take action $a = 0$.

The sender now wishes to find the lowest q subject to $1/(2-q) \geq r$, which yields $q^* = 2 - \frac{1}{r}$. Thus, stochastic upper censorship with $\omega^* = 0$ and $q^* = 2 - \frac{1}{r}$ is optimal. The probability that the receiver accepts the proposal is exactly equal to the probability that the signal realization is "accept":

$$\Pr[\text{proposal is accepted}] = \Pr[\omega = 1] + \Pr[\omega = 0] \cdot (1 - q^*) = \frac{1}{2} + \frac{1}{2} \cdot \frac{1-r}{r} = \frac{1}{2r}.$$

We now summarize the insights obtained in this example. If the state is continuous, then an optimal signal is deterministic upper censorship. If the state is discrete, then an optimal monotone partition is deterministic upper censorship, and an optimal signal is stochastic upper censorship.[10] In the remainder of the paper we will show that these insights extend to the case where the receiver's type is uncertain.

3. Continuous State

In this section, we provide necessary and sufficient conditions for the optimality of upper censorship when the state is continuous.

Let m be the expected state conditional on observing a realization of a signal s. Since the sender and receiver's utilities are linear in ω, they depend on the information about ω revealed by signal s only through the expected state m. In particular, the receiver chooses $a = 1$ if and only if $r \leq m$.

Let $V(m)$ denote the *indirect utility* of the sender conditional on m,

$$V(m) = \int_{r \leq m} \mathbb{E}\left[v(\omega, r) \middle| m \right] g(r) \mathrm{d}r = \int_0^m v(m, r) g(r) \mathrm{d}r, \quad m \in [0, 1]. \tag{2}$$

where $\mathbb{E}\left[v(\omega, r) \middle| m \right] = v(m, r)$ by the linearity of v in ω.

A function V is said to be *S-shaped* if it is convex below some threshold and concave above that threshold, or, equivalently, if V'' is single-crossing from above:

$$\text{there exists } \tau \text{ such that } \ V''(m) \geq (\leq)\, 0 \ \text{ for all } \ m < (>)\, \tau.$$

We now provide the criterion for the optimality of upper censorship. Recall that $m^* = \mathbb{E}[\omega | \omega \geq \omega^*]$.

Theorem 1. *Let V be S-shaped. Then, and only then, for all continuous F, an optimal signal is deterministic upper censorship whose cutoff state $\omega^* \in [0, \tau]$ satisfies:*

$$V(m^*) + V'(m^*)(m - m^*) \geq (\leq)\, V(m) \ \text{ for all } m \geq (<)\, \omega^*. \tag{3}$$

The criterion for the optimality of lower censorship is analogous.

[10] See, for example, Kolotilin (2015) for a detailed treatment of the case where there is no uncertainty about the receiver's type.

Remark 1. Let $-V$ be S-shaped. Then, and only then, for all continuous F, an optimal signal is deterministic lower censorship whose cutoff state $\omega^* \in [\tau, 1]$ satisfies:

$$V(m^*) + V'(m^*)(m - m^*) \geq (\leq)\, V(m) \text{ for all } m \leq (>)\, \omega^*.$$

Theorem 1 states that optimal persuasion takes a simple form of upper censorship when V is S-shaped.[11] Moreover, the theorem establishes that this result is tight, in the sense that if V is not S-shaped, then there exists a distribution F of the state such that no upper-censorship signal is optimal.

Note that every upper-censorship signal is a monotone partition. Therefore, Theorem 1 applies in both Scenario A, where the sender is free to choose any signal, and Scenario B, where the sender is constrained to choose a monotone partition.

The intuition behind Theorem 1 is as follows. Observe that when no information about the state is revealed, the receiver's best-response action does not change with the state. The more information is revealed, the more variable the receiver's behavior will be in response to this information. Consider an interval where V is concave. The sender would prefer not to reveal any information when the state is in this interval, since a certain outcome is preferred to any lottery with the same expected state. Conversely, consider an interval where V is convex. The sender would prefer to fully reveal the state in this interval, since now lotteries are preferred. If V is S-shaped, that is, it is convex below some threshold and concave above that threshold, the induced optimal persuasion takes the form of upper censorship.

When V is S-shaped, the sender's optimization problem is reduced to finding an optimal censorship cutoff ω^*. If the realized state ω is below the cutoff, then it is revealed to the receiver, so the expected utility of the sender is $V(\omega)$. If the realized state ω is above the cutoff, then the posterior expected state is $m^* = \mathbb{E}[\omega|\omega \geq \omega^*]$, so the expected utility of the sender conditional on $\omega \geq \omega^*$ is $V(m^*)$. The sender thus needs to solve the problem

$$\max_{\omega^* \in [0,1]} \int_0^{\omega^*} V(\omega) f(\omega) \mathrm{d}\omega + \int_{\omega^*}^1 V(m^*) f(\omega) \mathrm{d}\omega. \tag{4}$$

The expression in (3) represents the first-order condition to this problem and captures three possible cases: the boundary solutions $\omega^* = 0$ and $\omega^* = 1$ if the expression in (3) has the same sign for all ω, and an interior solution ω^* such that

$$V(m^*) + V'(m^*)(\omega^* - m^*) = V(\omega^*). \tag{5}$$

This first-order condition (5) is illustrated by Figure 1. The solid line is $V(m)$, and the dashed line is $V(m^*) + V'(m^*)(m - m^*)$, which is tangent to V at m^*.

[11] For a fixed distribution F, Proposition 3 in Kolotilin (2018) implies that upper censorship with cutoff ω^* is optimal if and only if (3) holds and $V(m)$ is convex on $[0, \omega^*]$. This result can be used to show that upper censorship is optimal when V is S-shaped. For completeness, we include a simple, self-contained proof, inspired by the proof of Theorem 1 in Dworczak and Martini (2019).

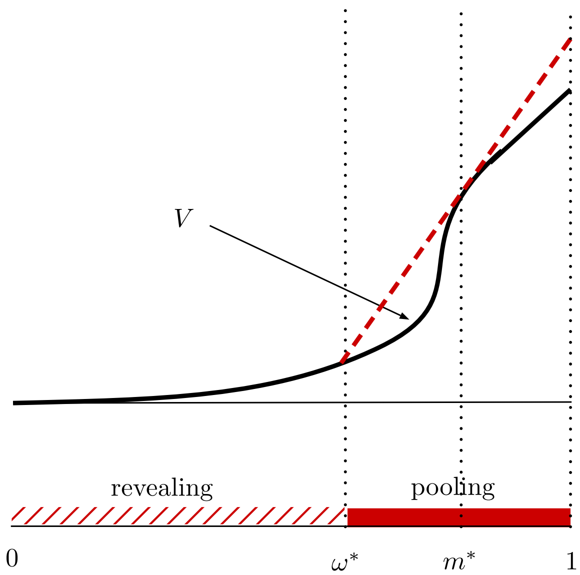

FIGURE 1. Optimal upper censorship with cutoff ω^*.

The solution of the sender's problem (4) is particularly simple if V is either globally convex or globally concave. In this case, the first-order condition in (3) has a constant sign, so either $\omega^* = 0$ or $\omega^* = 1$ must be optimal. If V is convex, then $\omega^* = 1$ is optimal, which corresponds to the fully informative signal. Similarly, if V is concave, then $\omega^* = 0$ is optimal, which corresponds to the completely uninformative signal. This is summarized in the following corollary.

Corollary 1. *An optimal signal is*

(i) *fully informative for all F if and only if V is convex;*

(ii) *completely uninformative for all F if and only if V is concave.*

3.1. **Canonical Case.** We now consider the canonical case where the sender's utility satisfies assumption (A_1).

The density g of receiver types r is said to be *log-concave* if $\ln g(r)$ is concave in r.[12] Note that $\ln g(r)$ is well defined in the canonical case.

In this case, the shape of the sender's indirect utility V defined by (9) is connected to the shape of the density g of receiver types as follows.

Lemma 1. *In the canonical case, V is S-shaped for all ρ if and only if g is log-concave.*

[12] Table 1 in Bagnoli and Bergstrom (2005) reports distributions with log-concave densities.

That is, if the density of receiver types g is log-concave, then the sender's indirect utility V is S-shaped. Moreover, this result is tight in the sense that if g is not log-concave, then there exists ρ such that V is not S-shaped.

By Theorem 1 and Lemma 1, we obtain a criterion for the optimality of upper censorship with the condition on the primitive of the model, the density g.

Theorem 2. *Consider the canonical case and let the density g be log-concave. Then, and only then, an optimal signal is deterministic upper censorship for all F and all ρ.*

The symmetric statement is also true: an optimal signal is *lower censorship* for all F and all ρ if and only if $-g$ is log-concave. Consequently, if both g and $-g$ are log-concave (that is, g is exponential), then there exists an optimal signal which is both upper censorship and lower censorship. There are only two signals with this property: fully informative and completely uninformative. This allows us to obtain conditions on the distribution of receiver types under which the optimal signal is polarized between fully informative and completely uninformative signals, as, for example, in Lewis and Sappington (1994) and Johnson and Myatt (2006).[13]

Corollary 2. *In the canonical case, an optimal signal is either fully informative or completely uninformative for all F and all ρ if and only if there exist $\lambda \in \mathbb{R}$ and $c > 0$ such that $g(r) = ce^{-\lambda r}$ for $r \in [0, 1]$.*

If $g(r) = ce^{-\lambda r}$, then the fully informative signal is optimal whenever $\rho \geq \lambda$ and the completely uninformative signal is optimal whenever $\rho \leq \lambda$ (and any signal is optimal when $\rho = \lambda$). In particular, if $\rho = 0$, the optimal signal is fully determined by the sign of λ, which is in turn determined by whether the mean of r is smaller or greater than $\frac{1}{2}$.

3.2. **Comparative Statics.** Theorem 2 allows for a sharp comparative statics analysis on the informativeness of the optimal signal.

We compare signals by their Blackwell informativeness (Blackwell, 1953). To compare upper-censorship signals s_1 and s_2, we only need to compare their cutoffs ω_1^* and ω_2^*. Signal s_1 is more informative than signal s_2 if $\omega_1^* \geq \omega_2^*$. Indeed, state $\omega \in [0, \omega_2^*)$ is fully revealed by both s_1 and s_2, and state $\omega \in [\omega_2^*, 1]$ is partially revealed by s_1 and not revealed at all by s_2, so s_1 is more informative than s_2.

[13] In the case of log-concave density g, the sets of upper- and lower-censorship signals are totally ordered by the rotation order of Johnson and Myatt (2006). If we restrict attention to the set of lower-censorship signals (which are generally suboptimal under log-concave g), then the rotation point is decreasing with respect to the rotation order. By Lemma 1 of Johnson and Myatt (2006), the sender's expected utility is quasiconvex and one of the extreme lower-censorship signals (the fully informative or completely uninformative signal) is optimal for the sender. But, if we restrict attention to the set of upper-censorship signals (which are optimal under log-concave g), then the rotation point is not decreasing with respect to the rotation order. Therefore, Lemma 1 of Johnson and Myatt (2006) does not apply, and an interior upper-censorship signal is generally optimal for the sender.

For the purpose of comparison, we extend the definition of distribution function G to the real line and assume that its density g is log-concave. Consider a family of distributions $G_{t,\sigma}$ of receiver types

$$G_{t,\sigma}(r) = G\left(\tau - t + \frac{r-\tau}{\sigma}\right),$$

where τ is the maximum point of g, and $t \in \mathbb{R}$ and $\sigma > 0$ are parameters. Let $g_{t,\sigma}(r)$ denote the corresponding density. Note that $g_{t,\sigma}(r)$ is log-concave for all $t \in \mathbb{R}$ and all $\sigma > 0$.

Because $g_{t,\sigma}$ is log-concave on $[0,1]$, deterministic upper censorship is optimal by Theorem 2. Let $\omega^*(\rho, t, \sigma) \in [0,1]$ be the optimal censorship cutoff as given by (3).

The *shift parameter* t shifts the distribution along the horizontal axis, and the *stretch parameter* σ stretches the distribution horizontally and symmetrically around the mode τ, so the limit of $\sigma \to 0$ leads to the unit mass on τ. We now show that the sender optimally discloses more information when:

(i) the sender's preferences are better aligned with the receiver's preferences (the alignment parameter ρ is greater),

(ii) the receiver is more reluctant to accept the proposal (the shift parameter t is greater),

(iii) the sender is more uncertain about the receiver's type (the stretch parameter σ is greater), provided $\rho \geq 0$.[14]

Theorem 3. *In the canonical case, a cutoff state of an optimal signal satisfies:*

(i) $\omega^*(\rho, t, \sigma)$ *is increasing in* ρ*;*

(ii) $\omega^*(\rho, t, \sigma)$ *is increasing in* t*;*

(iii) $\omega^*(\rho, t, \sigma)$ *is increasing in* σ *if* $\rho \geq 0$*.*

The intuition for part (i) is that for a higher ρ, the sender puts more weight on the receiver's utility, so she optimally endows the receiver with a higher utility by providing more information.

The intuition for part (ii) is that for a higher t, each type of the receiver has a greater cost of accepting the proposal, so to persuade the same type of the receiver, the sender needs to increase $\mathbb{E}[\omega|\omega \geq \omega^*]$ by expanding the full disclosure interval $[0, \omega^*]$.

The intuition for part (iii) is that for a higher σ, receiver types are more spread out, so to persuade the same mass of types, the sender again needs to increase $\mathbb{E}[\omega|\omega \geq \omega^*]$.

[14] The case $\rho \geq 0$ is the practically relevant case where the sender's utility is a weighted average of the receiver's action and utility: $\frac{1}{1+\rho} v(\omega, r) = \frac{1}{1+\rho} + \frac{\rho}{1+\rho}(\omega - r)$.

4. Discrete State

In this section we assume that the state is a discrete random variable.

Theorem 1′. *Let V be S-shaped. Then, and only then, for all discrete F,*

(A) *an optimal signal is stochastic upper censorship whose cutoff pair (ω^*, q^*) satisfies condition* (3)*;*

(B) *an optimal monotone partition is deterministic upper censorship whose cutoff pair (ω_d^*, q_d^*) satisfies $\omega_d^* = \omega^*$ and $q_d^* \in \{0, 1\}$.*

In the canonical case, we obtain the result analogous to Theorem 2.

Theorem 2′. *Consider the canonical case and let the density g be log-concave. Then, and only then, for all discrete F and all ρ,*

(A) *an optimal signal is stochastic upper censorship;*

(B) *an optimal monotone partition is deterministic upper censorship.*

Let us discuss how the comparative statics result (Theorem 3) changes. In Section 3.2, we argue that upper-censorship signals can be ordered by their censorship cutoffs. But when the state is discrete, a censorship cutoff is described by a pair (ω^*, q^*) where ω^* is a cutoff state and q^* is the probability of revealing this cutoff state when it realizes.

Consider two stochastic upper-censorship signals s_1 and s_2 with cutoff pairs (ω_1^*, q_1^*) and (ω_2^*, q_2^*). Denote $(\omega_1^*, q_1^*) \succeq (\omega_2^*, q_2^*)$ if $\omega_1^* > \omega_2^*$, or $\omega_1^* = \omega_2^*$ and $q_1^* \geq q_2^*$. Observe that s_1 is more (Blackwell) informative than s_2 if and only if $(\omega_1^*, q_1^*) \succeq (\omega_2^*, q_2^*)$. This comparison also applies to deterministic upper-censorship signals, with the constraint that q_1^* and q_2^* are in $\{0, 1\}$.

With the above order on cutoff pairs, Theorem 3 extend as follows:

Theorem 3′. *Consider the canonical case.*

(A) *A cutoff pair of an optimal signal satisfies:*

- (i) $(\omega^*(\rho, t, \sigma), q^*(\rho, t, \sigma))$ *is increasing in ρ;*
- (ii) $(\omega^*(\rho, t, \sigma), q^*(\rho, t, \sigma))$ *is increasing in t;*
- (iii) $(\omega^*(\rho, t, \sigma), q^*(\rho, t, \sigma))$ *is increasing in σ if $\rho \geq 0$.*

(B) *A cutoff pair of an optimal monotone partition satisfies:*[15]

- (i) $(\omega_d^*(\rho, t, \sigma), q_d^*(\rho, t, \sigma))$ *is increasing in ρ;*
- (ii) $(\omega_d^*(\rho, t, \sigma), q_d^*(\rho, t, \sigma))$ *is increasing in t.*

[15] Part (iii) also holds for an optimal monotone partition when, for example, the discrete support of the state ω is a regular grid with sufficiently many points. However, it is possible to construct a counterexample where part (iii) fails for an optimal monotone partition.

That is, the sender optimally discloses more information when she is less biased relative to the receiver (the alignment parameter ρ is greater), when the receiver is more reluctant to accept the proposal (the shift parameter t is greater), and, under Scenario A, when the sender is more uncertain about the receiver's type (the stretch parameter σ is greater).

Remark 2. Theorem 1′ and Theorem 2′ show that upper censorship emerges as an optimal persuasion strategy under the same conditions (V is S-shaped in Theorem 1′ and g is log-concave in Theorem 2′), regardless of whether the sender can choose an arbitrary signal or is constrained to choose a monotone partition. It may be tempting to conjecture that if the sender was allowed to choose an arbitrary partition (not necessarily monotone), the same result would hold. This, however, is not true.

For illustration, suppose that the sender wishes to maximize the probability of $a = 1$, and the receiver has type $r = \frac{1}{2}$ with certainty.[16] Let the state ω take values in $\{0, \frac{1}{4}, 1\}$ with probabilities $(\frac{1}{6}, \frac{2}{3}, \frac{1}{6})$, so the prior mean is $\frac{1}{3}$.

An optimal signal is stochastic upper censorship with cutoff pair $(\omega^*, q^*) = (\frac{1}{4}, \frac{1}{2})$, which sends the pooling message with certainty at $\omega = 1$ and with probability $1 - q^* = \frac{1}{2}$ at $\omega = \frac{1}{4}$, and reveals the state otherwise. The pooling message induces the posterior mean $m = \frac{1}{2}$, leading to $a = 1$; all other messages induce posterior means below $\frac{1}{2}$, leading to $a = 0$. Under the optimal signal, action $a = 1$ is chosen with probability $\frac{2}{3} \cdot \frac{1}{2} + \frac{1}{6} = \frac{1}{2}$.

An optimal *monotone* partition is deterministic upper censorship that fully reveals the state, where only $\omega = 1$ leads to $a = 1$. Under this partition, action $a = 1$ is chosen with probability $\frac{1}{6}$.

An optimal partition pools the extreme states $\omega = 0$ and $\omega = 1$, and reveals the intermediate state $\omega = \frac{1}{4}$. The pooling message induces the posterior mean $m = \frac{1}{2}$, leading to $a = 1$. This optimal partition is nonmonotone. Under this partition, action $a = 1$ is chosen with probability $\frac{1}{6} + \frac{1}{6} = \frac{1}{3}$, which is strictly better than the probability of $\frac{1}{6}$ obtained under the optimal monotone partition.

5. Application to Media Censorship

In this section, we apply our results to the problem of media censorship by the government. In the modern world, people obtain information about the government's state through various media sources such as television, newspapers, and internet blogs. Without the media, most people would not know what policies and reforms the government pursues and how effective they are. Media outlets have different positions on the political spectrum and differ substantially in how they select and present facts to cover the same news. People choose their sources of information based

[16] In this case, $V(m) = 1$ if $m \geq \frac{1}{2}$ and $V(m) = 0$ if $m < \frac{1}{2}$. Observe that the graph of V is S-shaped, and it can be approximated by a continuously differentiable S-shaped function.

on their political ideology and socioeconomic status. This information is valuable for significant individual decisions on migration, investment, and voting, to name a few. Individuals may not fully internalize externalities that their decisions impose on the society. Likewise, the government may not have the society's best interest at heart. To further its goals, the government then wishes to influence individual decisions by manipulating the information through media. In autocracies and countries with weak checks and balances, the government has power to censor the media content.

The government's problem of media censorship can be represented as the persuasion problem in Section 2. We apply our results to provide conditions for the optimality of upper-censorship policies that censor all media outlets except the most pro-government ones. Furthermore, we interpret our comparative statics results.

5.1. **Setup.** There is a continuum of heterogeneous citizens indexed by $r \in [0,1]$ distributed with G. Each citizen chooses between $a = 0$ and $a = 1$. The utility of a citizen of type r is given by

$$u(\theta, r, a_r, \bar{a}) = (\theta - r)a_r + \kappa(\theta, r, \bar{a}),$$

where $a_r \in \{0,1\}$ denotes the citizen's own action, $\bar{a} = \int a_r \mathrm{d}G(r)$ denotes the aggregate action in the society, $\theta \in [0,1]$ captures an unobserved benefit from action 1 as compared to action 0, and κ captures the impact of the aggregate action $\bar{a}$ on the citizen's utility. The term $(\theta - r)a_r$ is a private surplus of a citizen of type r. The term $\kappa(\theta, r, \bar{a})$ is an externality, because for a citizen of type r it is optimal to ignore this term and choose $a_r = 1$ if and only if $\theta \geq r$.

There is a government which is concerned with a weighted average of the social utility and the government's intrinsic benefit from the aggregate action. For a given θ, the government's utility is given by

$$\int_0^1 \nu(\theta, r, a_r, \bar{a}) \mathrm{d}G(r) + \delta\gamma(\theta, \bar{a}).$$

The term ν captures a citizen's utility from the government's perspective. We allow ν to be different from u to reflect paternalistic or other concerns. The term γ captures the government's intrinsic benefit from the aggregate action. The parameter $\delta \geq 0$ captures the weight of the aggregate action in the government's utility.

Let T be a distribution of the random variable θ. We assume that distributions G and T are independent and admit continuously differentiable and strictly positive densities. We also assume that κ, ν, and γ are linear in θ and continuously differentiable in r and $\bar{a}$. Furthermore, to simplify interpretations, we assume that ν and γ are non-decreasing in θ and $\bar{a}$, so that for the government, a high θ is a good news, and a higher aggregate action is preferable.

Citizens obtain information about the unobservable benefit θ through media outlets. Each media outlet is identified by its editorial policy $c \in [0,1]$, and it endorses action

$a = 1$ if $\theta \geq c$ and criticizes it if $\theta < c$.[17] A set C of media outlets is a finite subset of $[0, 1]$.[18]

The government's censorship policy is a set of the media outlets $X \subset C$ that are permitted to broadcast; so the rest of the media outlets are censored.

The timing is as follows. First, the government chooses a set $X \subset C$ of permitted media outlets. Second, state θ is realized, and each permitted media outlet endorses or criticizes action $a = 1$ according to its editorial policy. Finally, each citizen observes messages from all permitted media outlets, updates his beliefs about θ, and chooses an action.

5.2. **Discussion.** We now discuss interpretations of the key components of the media censorship application. As in Gehlbach and Sonin (2014), there can be various interpretations of the citizen's action $a = 1$, such as voting for the government, supporting a government's policy, or taking an individual decision that benefits the government. A citizen's type r can be interpreted as his ideological position or preference parameter: A citizen who is more supportive of the government has a smaller r.

A media outlet with a higher editorial policy c can be interpreted as being less loyal to the government because it criticizes the government on a larger set of states. An editorial policy $c \in C$ can therefore represent a slant or political bias of the outlet against the government and can be empirically measured as the frequency with which the outlet uses anti-government language. Gentzkow and Shapiro (2010) construct such a slant index for U.S. newspapers. Empirical findings of their paper suggest that the editorial policies of media outlets are driven by reader preferences, justifying our assumption of the existence of a large variety of editorial policies.[19] As in Suen (2004), Chan and Suen (2008), and Chiang and Knight (2011), the assumption of the binary media reports that communicate only whether the state θ is above some standard c can be justified by a cursory reader's preference for simple messages such as positive or negative opinions and yes or no recommendations.

The government's censorship of media outlets can take various forms. For example, the government can ban access to internet sites, withdraw licenses, disrupt financing, confiscate print materials and equipment, discredit media providers, and even arrest editors and journalists using broadly formulated legislation on combating extremism. In some countries, the government can exercise direct control over media editorial policies either through state ownership or administrative pressure.

[17]The tie-breaking in the event of $\theta = c$ is unimportant, as θ is a continuous random variable.

[18]We discuss the case of a continuum of media outlets in Section 5.5.

[19]Theoretical literature has explored the determinants of media slant of an outlet driven by its citizens (Mullainathan and Shleifer, 2005, Gentzkow and Shapiro, 2006, and Chan and Suen, 2008) and its owners (Baron, 2006, and Besley and Prat, 2006).

5.3. **Formulating Media Censorship as Persuasion.** We now show that the media censorship problem can be formulated as a linear persuasion problem, in which the government is a sender and a representative citizen is a receiver.

In our model, the citizens' and the government's utilities are linear in θ. Therefore, given any information from the media outlets, the utilities depend only on the posterior mean of θ.

Consider an arbitrary posterior mean of θ, denoted by $m \in [0,1]$. Each citizen of type r chooses $a_r = 1$ if and only if $r \leq m$. The externality term κ plays no role in this decision. Therefore, the aggregate action $\bar{a}$ is simply the mass of all citizens whose types do not exceed m, so $\bar{a} = G(m)$.

Next, using the citizens' optimal behavior, we derive the government's expected utility conditional on the posterior mean m:

$$\begin{aligned} V(m) &= \mathbb{E}\left[\int_0^1 \nu(\theta, r, a_r, G(m))\mathrm{d}G(r) + \delta\gamma(\theta, G(m))\,\middle|\, m\right] \\ &= \int_0^1 \nu(m, r, \mathbf{1}_{\{r\leq m\}}, G(m))\mathrm{d}G(r) + \delta\gamma(m, G(m)). \end{aligned} \tag{6}$$

This is the government's *indirect utility*. As in Section 2, the government now chooses a signal which is informative about the state to maximize its expected utility. However, in contrast to Section 2, here the government is restricted to signals that are implementable by a subset of a given set of media outlets.

Next, we show that, with an appropriate definition of the state, the restriction to signals implementable by a subset of a given set of media outlets can be formulated as a restriction to monotone partitional signals, as in Scenario B in Section 2. Thus we will be able to use our results in Section 4 to find the optimal censorship policies.

Recall that the information about θ is only available through the media outlets. Let ω denote a random variable equal to the conditional expectation of θ given the messages of all media outlets in C. Note that, as C is finite, ω is a discrete random variable with values in $[0,1]$. Let F denote the distribution of ω. From now on, we treat the random variable ω as the state, and denote by s a signal as defined in Section 2.

Let $\mathcal{H}_M$ denote the set of *all* distributions of the posterior mean state induced by all monotone partitional signals, and let $\mathcal{H}_C$ denote the set of *all* distributions of the posterior mean state induced by all finite subsets of the set C of media outlets. Observe that any set of media outlets induces a monotone partitional signal, that is, $\mathcal{H}_C \subset \mathcal{H}_M$. We now show that every outcome implementable by an arbitrary monotone partition in $\mathcal{H}_M$ is also implementable by a subset of media outlets.

Lemma 2. $\mathcal{H}_C = \mathcal{H}_M$.

For illustration, suppose that there is only one media outlet with the editorial policy $c \in (0,1)$. There are two observable events: $\theta \geq c$ and $\theta < c$. In each realized event, the citizens compute the posterior expectation of θ, which we will call ω. So, in this example, the distribution F of ω has two mass points, one for each of the two events.

Given these two mass points, every monotone partition of $[0,1]$ induces one of two possible distributions of the posterior mean state, (i) the mass points revealed, and (ii) the mass points pooled. Observe, however, that these distributions are implementable by permitting and censoring the media outlet, respectively.

A censorship policy is *upper censorship* if it censors all sufficiently disloyal media outlets. Specifically, there exists a cutoff $c^* \in C$ and an indicator $q^* \in \{0,1\}$ such that all media outlets whose editorial policies are below c^* are permitted, all media outlets whose editorial policies are above c^* are censored, and c^* is permitted if and only if $q^* = 1$. That is, $X = \{c \in C : c \le c^*\}$ if $q^* = 1$ and $X = \{c \in C : c < c^*\}$ if $q^* = 0$. Note that the *full censorship* policy $c^* = 0$ and $q^* = 0$ (where all media outlets are censored) and the *free media* policy $c^* = 1$ and $q^* = 1$ (where all media outlets are permitted) are the two extreme upper-censorship policies.

As shown above, the media censorship problem is equivalent to the linear persuasion problem in which the sender is restricted to monotone partitional signals and the sender's indirect utility V is given by (6). We thus apply Theorem 1′(B) to obtain the following result.

Theorem 1″. *If V is S-shaped, then an optimal censorship policy is upper censorship.*

To illustrate Theorem 1″, suppose the government is interested only in the aggregate action, so that $\nu = 0$ and γ depends only on the aggregate action $\bar{a}$; so $V(m) = \gamma(G(m))$. Thus, an upper-censorship policy is optimal if the composition function $\gamma(G(\cdot))$ is S-shaped. For example, this condition holds if the government is interested in reaching a certain approval threshold (e.g., a simple majority), so that γ is a step function. This condition also holds if γ is S-shaped and G is uniform or S-shaped with the same inflection point as γ.[20]

5.4. **Canonical Case.** We now impose more structure on the utilities to obtain a sharper result for the optimality of upper censorship and to perform a comparative statics analysis. Analogously to assumption (A_1), we assume that the government's utility is a weighted average of the citizens' utility and their aggregate action:

$$\int_0^1 u(\theta, r, a_r, \bar{a})\mathrm{d}G(r) + \delta\bar{a}, \qquad \text{where } u(\theta, r, a_r, \bar{a}) = (\theta - r)a_r + \zeta(r)\bar{a} \tag{A$_2$}$$

for some continuously differentiable function ζ. Define

$$\beta = \int_0^1 \zeta(r)\mathrm{d}G(r) + \delta.$$

[20] This condition holds in many special cases where γ and G are S-shaped even when γ and G have different inflection points.

The term $\int_0^1 \zeta(r)\mathrm{d}G(r)$ is the government's expected bias towards the citizens' action $a = 1$ due to the citizens' externality, and the term δ is the government's intrinsic bias towards a greater aggregate action $\bar{a}$. Thus, we interpret β as the government's aggregate bias. This decomposition of the bias allows for different interpretations why the government is biased. So, β can be high because the government is too self-serving (high δ), or because the government is benevolent (low δ) but wishes to internalize strong positive externalities of the citizens (high $\zeta(r)$). To ease interpretations, we assume that $\beta > 0$.[21]

Under Assumption (A_2), the government's indirect utility V given by (6) becomes

$$V(m) = \int_0^m (m - r)\mathrm{d}G(r) + \beta\bar{a} = \int_0^m v(m, r)\mathrm{d}G(r),$$

where

$$v(m, r) = \beta\left(1 + \frac{1}{\beta}(m - r)\right).$$

This is the same as v given by Assumption (A_1) with $\rho = \beta^{-1}$, up to rescaling by a positive constant. Consequently, we can apply Theorem $2'$(B) to obtain the following result.

Theorem $2''$. *Let (A_2) hold. If the density g of citizens types is log-concave, then an optimal censorship policy is upper censorship.*

We now apply the comparative statics analysis presented in Section 3.2. Note that the upper-censorship policies are ordered according to the amount of information transmitted to the citizens, in the sense of Blackwell (1953). A greater censorship threshold c^* means that more media outlets are permitted. With this order in mind, we apply Theorem $3'$(B) to make a comparative statics analysis on the amount of information that is optimally disclosed by the government.

First, the censorship cutoff is increasing in the alignment parameter ρ. Recall that this is a reciprocal of the government's bias, $\rho = (1 - \delta)/\beta$. This means that the government optimally discloses more information (the censorship cutoff is greater) when it is less biased (β is smaller). Intuitively, as β decreases, the government puts more weight on the citizens' utility, so it optimally endows the population with a higher utility by censoring fewer media outlets and disclosing more information.

Second, the censorship cutoff is increasing in the magnitude t of the horizontal shift of the density g. A greater t corresponds to a greater opportunity cost of action $a = 1$ for each citizen. This means that the government optimally discloses more information (the censorship cutoff is greater) when the citizens are more difficult to persuade to take action $a = 1$ (parameter t is greater). Informally speaking, to persuade the same type of the citizen, the government needs to increase the posterior

[21] If $\beta < 0$, then swapping the roles of $a = 0$ and $a = 1$ reverses the sign of the bias. If $\beta = 0$, then the government's utility and the citizens' private interests coincide, so it is trivially optimal to permit all media outlets.

mean. But the expected posterior mean must be equal to the prior mean, so it is not possible to increase all posterior means. Due to the log-concave shape of the density of the citizens' types, this tradeoff is resolved by increasing the posterior mean of the pooling interval, $\mathbb{E}[\omega|\omega \geq c^*]$. This is done by shrinking the pooling interval $[c^*, 1]$, that is, increasing the censorship cutoff c^*.

Third, when the number of media outlets is sufficiently large, the censorship cutoff is increasing in the stretch factor σ of the density g around its peak τ (see Footnote 15). A greater σ corresponds to a more dispersed distribution of the opportunity cost in the population. This means that the government optimally discloses more information (the censorship cutoff is greater) when the citizens are more diverse in how difficult they are to persuade. Intuitively, when the types are more spread out, persuading the same mass of types requires a greater posterior mean which leads to a higher censorship cutoff, similarly to the comparative statics in t discussed above.

5.5. **Extensions and Open Questions.** Let us now consider a few extensions of our model of media censorship.

In our model, the set of media outlets is exogenous, and the government's only instrument is censorship. We now consider three alternative ways of expanding the government's instruments of influence.

First, suppose that the government is able not only to censor existing media outlets, but also to introduce new media outlets with chosen editorial policies. This is equivalent to our censorship model where all media outlets in $[0, 1]$ are initially available, and the government can censor any subset of them. When all media outlets in $[0, 1]$ are permitted, the revealed information about θ (which we call ω) is θ itself. Thus, $\omega = \theta$ is a continuous random variable with distribution $F = T$. So, we can now apply our results for the continuous state from Section 3 instead of those for the discrete state in Section 4, reaching the same conclusion about the optimality of upper censorship.

Second, suppose that the government can garble information available from media outlets. That is, the government is not restricted to monotone partitions, it can create arbitrary signals about state ω for the citizens to observe. This becomes a general persuasion problem, and our Theorems $1'$(A), $2'$(A), and $3'$(A) apply.

Third, suppose that the government is able to restrict not only which media outlets are permitted, but also how many media outlets each citizen can choose to observe. In this extension, the citizens are not allowed to communicate with one another (otherwise they could share the information, thus observing all permitted media outlets indirectly). This extension does not affect our results, as long as each citizen is allowed to access at least one media outlet of his choice, as in Chan and Suen (2008). Intuitively, this is because each citizen categorizes the information from the media outlets into "good news" where $a = 1$ is optimal and "bad news" where $a = 0$ is optimal. Because the information from the media outlets induces a monotone partition, it means that "good news" is separated from "bad news" by a threshold media outlet that depends on the citizen's type. That is, it is sufficient to observe a single threshold media outlet to distinguish "good news" from "bad news".

In our model, each citizen's utility depends on the aggregate action $\bar{a}$ through the externality term $\kappa(\theta, r, \bar{a})$ which does not affect the chosen action. Let us relax this assumption, so that a citizen's optimal choice can depend on $\bar{a}$. We can still write the sender's indirect utility V as a function of the posterior mean state and apply our results. However, V is no longer uniquely determined by the primitives of the model. It is now endogenous and depends on an equilibrium the citizens play, as each citizen's optimal action now depends on what all citizens do in equilibrium. For example, given the same information about the state, a citizen could prefer to choose $a = 1$ if and only if sufficiently many citizens choose the same action, so $\bar{a}$ is large enough. This creates the problem of multiplicity of equilibria and, as a consequence, the dependence of optimal censorship on equilibrium selection.

As a side, our media censorship problem can also be applied to spatial voting models, as in Chiang and Knight (2011). Consider a government party ($p = G$) and an opposition party ($p = O$) competing in an election. If party p wins, a voter with an ideological position r gets utility $w_p - (r - r_p)^2$, where w_p is the quality or valence of party p, and r_p is the ideology or policy platform of party p. Voters know the parties' ideologies and obtain information about the parties' qualities from all available media outlets. Each voter supports the party that maximizes his expected utility. Within this context, our analysis still applies, because the voter's utility difference between the government and opposition parties is proportional to $\theta - r$, where $\theta = (w_G - w_O + r_O^2 - r_G^2)/2(r_O - r_G)$.

There are a few more extensions that can be relevant in applications. First, instead of complete censorship of a media outlet, there can be a cost of accessing it. For example, an international news channel can be censored by a local government, but citizens may still access it through VPN at some cost. Second, it can be costly for the government to censor media outlets. So, another important question to answer is how a government may prioritize censoring. Finally, citizens can incur some cost of following each media outlet. While we have already mentioned that citizens gain no benefit from following more than one outlet, it is entirely possible for them to stop watching news altogether if it is sufficiently uninformative. These extensions are nontrivial and left for future research.

Appendix

Proof of Theorem 1. Suppose that V is S-shaped. For $m, \omega^* \in [0, 1]$ and $m^*(\omega^*) = \mathbb{E}[\omega|\omega \geq \omega^*]$, define

$$\overline{V}(m, \omega^*) = V(m^*(\omega^*)) + V'(m^*(\omega^*))(m - m^*(\omega^*)). \tag{7}$$

Since $V(m)$ is convex on $[0,\tau]$ and concave on $[\tau,1]$, there exists $\omega^* \in [0,\tau]$ that satisfies condition (3), that is,

$$\overline{V}(m,\omega^*) \geq (\leq) V(m) \text{ for all } m \geq (<)\, \omega^*. \tag{8}$$

This is true, because $\overline{V}(\omega^*,\omega^*) - V(\omega^*)$ is single-crossing from below. To see this, consider Figure 1. Observe that (3) can hold only if ω^* is on the convex part of V and m^* is on the concave part of V. As the censorship cutoff ω^* increases, the posterior mean state of the pooling message $m^*(\omega^*)$ also increases. But because V is concave at $m^*(\omega^*)$, the dashed tangent line, $\overline{V}(m,\omega^*)$, becomes flatter in m, so it crosses the solid line, $V(m)$, at a smaller m.

Consider an arbitrary signal s. Let H be the distribution of $m = \mathbb{E}[\omega|s]$ induced by signal s. Since each signal is a garbling of the fully informative signal, F is a mean-preserving spread of H. Thus, the sender's expected utility is smaller under signal s than under the upper-censorship signal with cutoff ω^* as follows from

$$\begin{aligned}
\int_0^1 V(m)\mathrm{d}H(m) &\leq \int_0^1 \max\{V(m),\overline{V}(m,\omega^*)\}\mathrm{d}H(m) \\
&\leq \int_0^1 \max\{V(m),\overline{V}(m,\omega^*)\}\mathrm{d}F(m) \\
&= \int_0^{\omega^*} V(m)\mathrm{d}F(m) + \int_{\omega^*}^1 \overline{V}(m,\omega^*)\mathrm{d}F(m) \\
&= \int_0^{\omega^*} V(m)\mathrm{d}F(m) + \int_{\omega^*}^1 V(m^*(\omega^*))\mathrm{d}F(m),
\end{aligned}$$

where the second line holds because $\max\{V(m),\overline{V}(m,\omega^*)\}$ is convex in m and F is a mean-preserving spread of H, the third line holds by (8), and the last line holds by (7) and by $\int_{\omega^*}^1 (m - m^*(\omega^*))\mathrm{d}F(m) = 0$.

Conversely, suppose that V is not S-shaped. Then there exist $0 \leq m_1 < m_2 \leq m_3 < m_4 \leq 1$ such that $V''(m) < 0$ for $m \in (m_1,m_2)$, $V''(m) = 0$ for $m \in [m_2,m_3]$, and $V''(m) > 0$ for $m \in (m_3,m_4)$, because V'' is continuous by assumption. It is straightforward to show that there exists $\widehat{m}^* \in (m_1,m_2)$ (sufficiently close to m_2) and $\widehat{\omega}^* \in (m_3,m_4)$ such that

$$\begin{aligned}
V(\widehat{m}^*) + V'(\widehat{m}^*)(m-\widehat{m}^*) &< V(m) \quad \text{for all } m \in (\widehat{\omega}^*, m_4], \\
V(\widehat{m}^*) + V'(\widehat{m}^*)(m-\widehat{m}^*) &> V(m) \quad \text{for all } m \in [m_1, \widehat{\omega}^*] \setminus \{\widehat{m}^*\}.
\end{aligned}$$

Moreover, there exists a continuous distribution F such that the support of F is a subset of $[m_1,m_4]$ and $\widehat{m}^* = \mathbb{E}[\omega|\omega \leq \widehat{\omega}^*]$. Using a chain of inequalities similar to the one above, it is easy to show that lower-censorship signal $\widehat{s}^*$ with cutoff $\widehat{\omega}^*$ is optimal. Furthermore, it is uniquely optimal, as both inequalities hold with equalities only for signal $\widehat{s}^*$. This, in turn, implies that any upper-censorship signal is suboptimal. □

Proof of Lemma 1. Notice that, under assumption (A_1), we have

$$V''(m) = g'(m) + \rho g(m) \text{ for } m \in [0,1].$$

Recall that V is S-shaped if and only if $V''(m) = g'(m) + \rho g(m)$ is single-crossing from above. Using Proposition 1 in Quah and Strulovici (2012), it is easy to show that $V''(m)$ is single-crossing from above for all $\rho \in \mathbb{R}$ if and only if $g'(m)/g(m)$ is nonincreasing in m (that is, $\ln g(m)$ is concave). □

Proof of Theorem 2. Immediate by Theorem 1 and Lemma 1. □

Proof of Theorem 3. *Part (i).* We fix t and σ, and vary ρ. Without loss of generality, let $t = 0$ and $\sigma = 1$ (so that $G_{t,\sigma} = G$). Consider a signal that induces the distribution H of posterior mean m. This signal implements the receiver's interim utility U given by

$$U(r) = \int_r^1 (m - r)\mathrm{d}H(m) \text{ for } r \in [0,1],$$

which holds because the receiver acts if and only if $m \geq r$. Since types $r < 0$ always choose $a = 1$ and types $r > 1$ always choose $a = 0$, we can exclude them from consideration and assume that $\mathrm{supp}(G) = [0,1]$. The sender's expected utility is then

$$\begin{aligned}\int_0^1 V(m)\mathrm{d}H(m) &= \int_0^1 \int_0^m (1 + \rho(m - r))\mathrm{d}G(r)\mathrm{d}H(m)\\ &= \int_0^1 \int_r^1 (1 + \rho(m - r))\mathrm{d}H(m)\mathrm{d}G(r)\\ &= \int_0^1 (1 - H(r) + \rho U(r))\mathrm{d}G(r).\end{aligned}$$

Consider $\rho_2 > \rho_1$. Suppose by contradiction that the corresponding optimal upper-censorship signals s_2 and s_1 are such that s_1 is strictly more informative than s_2. Let U_2 and U_1 be the receiver's interim utilities implemented by s_2 and s_1. Since the sender prefers signal s_2 under ρ_2 and signal s_1 under ρ_1, we have

$$\begin{aligned}\int_0^1 (1 - H_2(r) + \rho_2 U_2(r))\,\mathrm{d}G(r) &\geq \int_0^1 (1 - H_1(r) + \rho_2 U_1(r))\,\mathrm{d}G(r),\\ \int_0^1 (1 - H_1(r) + \rho_1 U_1(r))\,\mathrm{d}G(r) &\geq \int_0^1 (1 - H_2(r) + \rho_1 U_2(r))\,\mathrm{d}G(r).\end{aligned}$$

Summing up these inequalities gives:

$$(\rho_2 - \rho_1)\int_0^1 (U_2(r) - U_1(r))\,\mathrm{d}G(r) \geq 0,$$

leading to a contradiction because $U_1(r) \geq U_2(r)$ for all r with strict inequality for some r given that s_1 is strictly more informative than s_2.

Part (ii). We fix $\rho \in \mathbb{R}$ and $\sigma = 1$, and vary t. The *indirect utility* of the sender conditional on $m \in [0,1]$ is given by

$$V_t(m) = \int_{r \le m} (1+\rho(m-r)) g_{t,1}(r) \mathrm{d}r = \int_{r \le m-t} (1+\rho(m-t-r)) g(r) \mathrm{d}r = V(m-t). \quad (9)$$

By Theorem 1, when $t_0 = 0$, an optimal cutoff ω_0^* satisfies

$$V(m) \le V(m_0^*) + V'(m_0^*)(m - m_0^*) \text{ for } m \in [\omega_0^*, 1].$$

When V is S-shaped and $t_1 < 0$, it is easy to see that

$$V(m - t_1) \le V(m_0^* - t_1) + V'(m_0^* - t_1)(m - m_0^*) \text{ for } m \in [\omega_0^*, 1],$$

which implies that, when $t = t_1$, the above inequality changes its sign at some ω_1^* that satisfies $\omega_1^* \le \omega_0^*$. It follows from (9) that ω_1^* is an optimal cutoff for $t = t_1$, and as we established above, $\omega_1^* \le \omega_0^*$.

Part (iii). We fix $\rho \ge 0$ and $t = 0$, and vary σ. The *indirect utility* of the sender conditional on $m \in [0,1]$ is given by

$$\begin{aligned} V_{\frac{\rho}{\sigma},\sigma}(m) &= \int_{r \le m} \left(1 + \frac{\rho}{\sigma}(m - r)\right) \mathrm{d}G_{0,\sigma}(r) \\ &= \int_{r \le m} \left(1 + \frac{\rho}{\sigma}(m - r)\right) g\left(\frac{r - (1-\sigma)\tau}{\sigma}\right) \frac{1}{\sigma} \mathrm{d}r \\ &= \int_{r \le \frac{m-(1-\sigma)\tau}{\sigma}} \left(1 + \rho\left(\frac{m - (1-\sigma)\tau}{\sigma} - r\right)\right) g(r) \mathrm{d}r \\ &= V_{\rho,1}\left(\frac{m - (1-\sigma)\tau}{\sigma}\right). \end{aligned} \quad (10)$$

Let $\sigma_0 = 1$, $\sigma_1 < 1$ and $\rho_1 = \rho/\sigma_1$. To compare the optimal cutoffs under parameters (ρ, σ_0) and (ρ, σ_1), we consider two steps, first comparing the cutoffs between (ρ, σ_0) and (ρ_1, σ_1), and then comparing the cutoffs between (ρ_1, σ_1) and (ρ, σ_1).

First, we compare an optimal cutoff ω_0^* under $(\rho, \sigma_0) = (\rho, 1)$ and an optimal cutoff ω_1^* under (ρ_1, σ_1). By Theorem 1, ω_0^* satisfies

$$V_{\rho,1}(m) \le V_{\rho,1}(m_0^*) + V'_{\rho,1}(m_0^*)(m - m_0^*) \text{ for } m \in [\omega_0^*, 1].$$

When V is increasing (implied by $\rho \ge 0$) and S-shaped, and $\sigma_1 < 1$, it is easy to see that, for all $m \in [\omega_0^*, 1]$,

$$V_{\rho,1}\left(\frac{m - (1-\sigma_1)\tau}{\sigma_1}\right) \le V_{\rho,1}\left(\frac{m_0^* - (1-\sigma_1)\tau}{\sigma_1}\right) + V'_{\rho,1}\left(\frac{m_0^* - (1-\sigma_1)\tau}{\sigma_1}\right)(m - m_0^*),$$

which implies that, when $\sigma = \sigma_1 < 1$, the above inequality changes its sign at some ω_1^* that satisfies $\omega_1^* \le \omega_0^*$. It follows from (10) that ω_1^* is an optimal cutoff under $(\rho/\sigma, \sigma_1) = (\rho_1, \sigma_1)$, and as we established above, $\omega_1^* \le \omega_0^*$.

Second, we compare the optimal cutoff ω_1^* under (ρ_1, σ_1) and the optimal cutoff ω_2^* under (ρ, σ_1). Because $\rho < \rho_1 = \rho/\sigma_1$, by part (i) we have $\omega_2^* \leq \omega_1^*$.

Combining $\omega_1^* \leq \omega_0^*$ and $\omega_2^* \leq \omega_1^*$, we obtain $\omega_2^* \leq \omega_0^*$. □

Proof of Theorem 1′. *Part (A).* The proof is analogous to that of Theorem 1 and is thus omitted.

Part (B). The *only if* statement follows from that in part (A) (see the proof of Theorem 1) and the fact that we can always ensure that an optimal stochastic censorship (ω^*, q^*) is deterministic (so $q^* \in \{0, 1\}$) by the choice of the support and distribution of ω. Thus, we only need to prove the *if* statement.

Let F be a distribution of ω with support on n values, $0 < \omega_0 < ... < \omega_{n-1} < 1$. This is without loss of generality, because if 0 and/or 1 were in the support of F, we could redefine the state as $\tilde{\omega} = (\omega + \varepsilon)/(1 + 2\varepsilon)$ and the type as $\tilde{r} = (r + \varepsilon)/(1 + 2\varepsilon)$.

Let $X \subset \{\omega_1, ..., \omega_{n-1}\}$ identify a set of cutoffs defining a monotone partition: for each $\omega_i \in X$, the receiver is informed whether $\omega < \omega_i$ or $\omega \geq \omega_i$. Note that we exclude ω_0 from the set of cutoffs, because $\omega \geq \omega_0$ always holds.

Note that X is deterministic upper censorship if and only if there exists a censorship threshold $\omega^* \in [0, 1]$ such that, for each $i = 1, ..., n-1$, $\omega_i \in X$ if and only if $\omega_i < \omega^*$, that is, we can assume that $q^* = 0$ without loss of generality.

Let V_X be the sender's expected utility when the monotone partition is given by X.

If $n = 2$, then there are only two partitions: $X = \{\omega_1\}$ (full disclosure) and $X = \varnothing$ (no disclosure). Both are upper censorships, so part (B) holds trivially.

If $n = 3$, then there are four monotone partitions: $X = \{\omega_1, \omega_2\}$ (full disclosure), $X = \{\omega_2\}$, $X = \{\omega_1\}$, and $X = \varnothing$ (no disclosure). Only $X = \{\omega_2\}$ is not upper censorship. The next lemma proves that $X = \{\omega_2\}$ cannot be uniquely optimal when V is S-shaped, thus proving the optimality of upper censorship for $n = 3$.

Denote

$$m_{jk} = \mathbb{E}\big[\omega | \omega \in [\omega_j, \omega_k]\big].$$

For convenience of notation, we write $m_j = \omega_j$.

Lemma 3. *Let V be S-shaped with an inflexion point τ. Let ω be a discrete random variable with support on $\{m_0, m_1, m_2\}$ and probabilities p_0, p_1, p_2, where $0 < m_0 < m_1 < m_2 < 1$. Then:*

(i) if $\tau \leq m_0$, then $V_{\{m_2\}} \leq V_\varnothing$;

(ii) if $\tau \geq m_1$, then $V_{\{m_2\}} \leq V_{\{m_1, m_2\}}$;

(iii) if $m_0 < \tau < m_1$, then $V_{\{m_2\}} \leq \max\{V_\varnothing, V_{\{m_1, m_2\}}\}$.

Proof. Define $p_{01} = p_0 + p_1$ and $p_{02} = p_{01} + p_2 = 1$. Observe that

$$m_{01} = \frac{p_0 m_0 + p_1 m_1}{p_0 + p_1} \quad \text{and} \quad m_{02} = \frac{p_0 m_0 + p_1 m_1 + p_2 m_2}{p_0 + p_1 + p_2}.$$

We thus have

$$\begin{aligned} V_\varnothing &= p_{02} V(m_{02}) = (p_{01} + p_2) V\left(\frac{p_{01} m_{01} + p_2 m_2}{p_{01} + p_2}\right), \\ V_{\{m_2\}} &= p_{01} V(m_{01}) + p_2 V(m_2) = (p_0 + p_1) V\left(\frac{p_0 m_0 + p_1 m_1}{p_0 + p_1}\right) + p_2 V(m_2), \\ V_{\{m_1, m_2\}} &= p_0 V(m_0) + p_1 V(m_1) + p_2 V(m_2). \end{aligned}$$

Clearly,

$$m_0 < m_{01} < m_1 \quad \text{and} \quad m_0 < m_{02} < m_2.$$

Therefore,

$$V_{\{m_2\}} \leq V_\varnothing \iff$$

$$\frac{p_{01}}{p_{01} + p_2} V(m_{01}) + \frac{p_2}{p_{01} + p_2} V(m_2) \leq V\left(\frac{p_{01}}{p_{01} + p_2} m_{01} + \frac{p_2}{p_{01} + p_2} m_2\right) = V(m_{02}), \tag{11}$$

and

$$V_{\{m_2\}} \leq V_{\{m_1, m_2\}} \iff$$

$$V\left(\frac{p_0}{p_0 + p_1} m_0 + \frac{p_1}{p_0 + p_1} m_1\right) \leq \frac{p_0}{p_0 + p_1} V(m_0) + \frac{p_1}{p_0 + p_1} V(m_1). \tag{12}$$

Suppose that $\tau \leq m_{01}$, so V is concave on $[m_{01}, m_2]$. Then, (11) holds by Jensen's inequality. In particular, it holds for $\tau \leq m_0 < m_{01}$, thus proving part (i).

Suppose that $\tau \geq m_1$, so V is convex on $[m_0, m_1]$. Then, (12) holds by Jensen's inequality, thus proving part (ii).

Finally, suppose that $m_{01} < \tau < m_1$. Let $(\tau, V(\tau))$ lie *below (or on)* the straight line A connecting $(m_0, V(m_0))$ and $(m_2, V(m_2))$, as illustrated on Figure 2. By convexity of $V(m)$ for $m < \tau$, it must be the case that $V(m_{01})$ lies below (or on) the dashed line B. By concavity of $V(m)$ for $m > \tau$, it must be the case that $V(m_1)$ lies above (or on) the dashed line C. Consequently, (12) is satisfied.

Alternatively, let $(\tau, V(\tau))$ lie *above* the straight line A connecting $(m_0, V(m_0))$ and $(m_2, V(m_2))$, as illustrated on Figure 3. If $V_{\{m_2\}} \leq V_{\{m_1, m_2\}}$, then the proof of part (iii) is complete. So, suppose that $V_{\{m_2\}} > V_{\{m_1, m_2\}}$, so (12) is violated.

By convexity of $V(m)$ for $m < \tau$, it must be the case that $V(m_{01})$ lies below (or on) the line B. Since (12) is violated, $V(m_{01})$ must lie between the lines A and B. Recall that $m_{01} < m_{02} < m_2$. Observe that $V(m_{02})$ must be in the shaded area on Figure 3. Indeed, if $m_{02} < \tau$, then by convexity of $V(m)$ for $m < \tau$, it must be the case that $V(m_{02})$ lies between (or on) the lines B and D. If $m_{02} \geq \tau$, then by concavity of

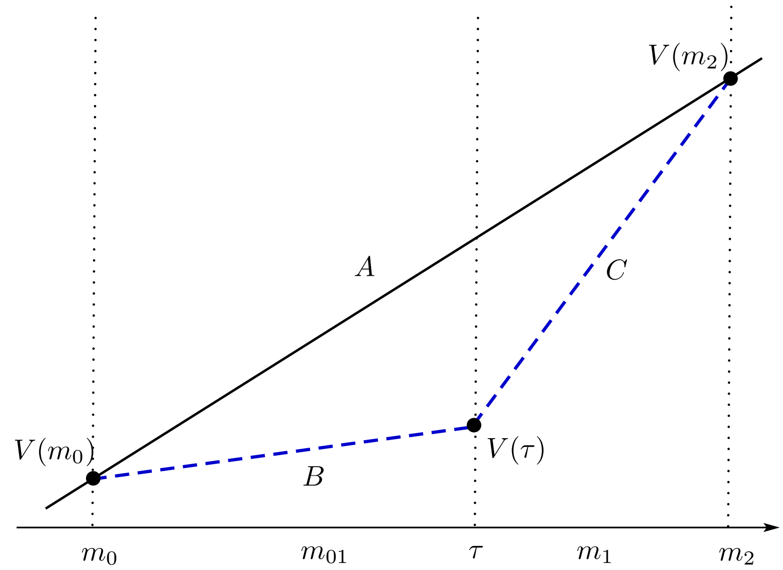

FIGURE 2. The case where $(\tau, V(\tau))$ lies below (or on) the straight line connecting $(m_0, V(m_0))$ and $(m_2, V(m_2))$.

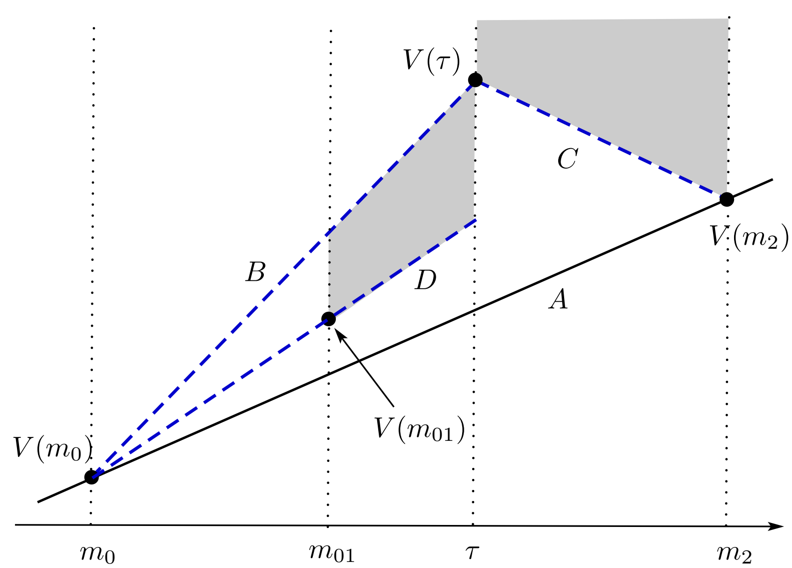

FIGURE 3. The case where $(\tau, V(\tau))$ lies above the straight line connecting $(m_0, V(m_0))$ and $(m_2, V(m_2))$.

$V(m)$ for $m > \tau$, it must be the case that $V(m_{02})$ lies above (or on) line C. It follows that (11) is satisfied. $\square$

We now adapt Lemma 3 to prove the optimality of upper censorship for $n \geq 4$.

Fix a monotone partition $X \subset \{\omega_1, ..., \omega_{n-1}\}$ that is not upper censorship, that is, there exists $i \in \{2, ..., n-1\}$ such that

$$\omega_{i-1} \notin X \quad \text{and} \quad \omega_i \in X. \tag{13}$$

Denote $\omega_n = 1$. Define j and k as follows. If $X \cap \{\omega_1, ..., \omega_{i-2}\} \neq \varnothing$, let ω_j be the largest element in X smaller than i; otherwise let $\omega_j = \omega_0$. If $X \cap \{\omega_{i+1}, ..., \omega_{n-1}\} \neq \varnothing$, let ω_k be the smallest element in X greater than i; otherwise let $\omega_k = \omega_n = 1$. Define

$$m_0 = \omega_j, \quad m_1 = \omega_{i-1}, \quad m_2 = \omega_i,$$

and

$$p_0 = \frac{\Pr[\omega \in [\omega_j, \omega_{i-1})]}{\Pr[\omega \in [\omega_j, \omega_k)]}, \quad p_1 = \frac{\Pr[\omega \in [\omega_{i-1}, \omega_i)]}{\Pr[\omega \in [\omega_j, \omega_k)]}, \quad p_2 = \frac{\Pr[\omega \in [\omega_i, \omega_k)]}{\Pr[\omega \in [\omega_j, \omega_k)]}.$$

To apply Lemma 3, we identify $X \backslash \{\omega_i\}$ with $\varnothing$, X with $\{m_2\}$, and $X \cup \{\omega_{i-1}\}$ with $\{m_1, m_2\}$. By Lemma 3, $V_X \leq \max\{V_{X\backslash\{\omega_i\}}, V_{X\cup\{\omega_{i-1}\}}\}$. Thus, replacing X with $X\backslash\{\omega_i\}$ if $V_X \leq V_{X\backslash\{\omega_i\}}$ and with $X \cup \{\omega_{i-1}\}$ otherwise weakly improves the sender's expected utility. Repeatedly applying such replacements yields upper censorship. Indeed, this procedure terminates after a finite number of replacements, because, by Lemma 3, for $\omega_j \geq \tau$, we replace X with $X\backslash\{\omega_i\}$ and, for $\omega_{i-1} \leq \tau$, we replace X with $X \cup \{\omega_{i-1}\}$.

We thus have shown that when V is S-shaped, when optimizing over the set of monotone partitions, we can restrict attention to deterministic upper censorship. It remains to show that an optimal deterministic censorship cutoff ω_d^* coincides with an optimal stochastic censorship cutoff ω^* from part (A). To see this, it is straightforward to show that

$$V^*(\omega^*, q^*) = \int_{[0,\omega^*)} V(\omega)\mathrm{d}F(\omega) + V(\omega^*)q^* \Pr[\omega = \omega^*] \\ + V(m^*(\omega^*, q^*))((1 - q^*) \Pr[\omega = \omega^*] + \Pr[\omega > \omega^*]) \tag{14}$$

is single-peaked in (ω^*, q^*) in the Blackwell informativeness order. Since the two deterministic censorship cutoff pairs closest to (ω^*, q^*) in the Blackwell informativeness order are $(\omega^*, 0)$ and $(\omega^*, 1)$, we obtain $\omega_d^* = \omega^*$. □

Proof of Theorem 2′. Immediate by Theorem 1′ and Lemma 1. □

Proof of Theorem 3′. The proof for an optimal signal is analogous to that of Theorem 3 and is thus omitted. The same applies to part (i) for an optimal monotone partition. Thus, it remains to prove part (ii) for an optimal monotone partition.

We fix ρ and $\sigma = 1$, and vary t continuously. If $\omega^* \notin \mathrm{supp}(F)$, then an optimal monotone partition is an optimal signal, because all $q^* \in [0, 1]$ are optimal. Thus, part (ii) in (B) holds by part (ii) in (A). If $\omega^* \in \mathrm{supp}(F)$ and $q^* \in (0, 1)$, then either $(\omega_d^*, q_d^*) = (\omega^*, 0)$ or $(\omega_d^*, q_d^*) = (\omega^*, 1)$ is optimal, because the sender's expected utility V^* given by (14) is single-peaked in (ω^*, q^*) in the Blackwell informativeness order.

Therefore, it suffices to consider values of t at which the sender is indifferent between $(\omega^*, 0)$ and $(\omega^*, 1)$ and to show that the sender strictly prefers $(\omega^*, 1)$ to $(\omega^*, 0)$ at a slightly higher t:

$$V_t^*(\omega^*, 0) = V_t^*(\omega^*, 1) \implies \frac{\mathrm{d}V_t^*(\omega^*, 0)}{\mathrm{d}t} \leq \frac{\mathrm{d}V_t^*(\omega^*, 1)}{\mathrm{d}t}. \tag{15}$$

The condition $V_t^*(\omega^*, 0) = V_t^*(\omega^*, 1)$ can be written as

$$V_t(m^*(\omega^*, 0)) \Pr[\omega \geq \omega^*] = V_t(\omega^*) \Pr[\omega = \omega^*] + V_t(m^*(\omega^*, 1)) \Pr[\omega > \omega^*]$$

$$\iff V_t(m^*(\omega^*, 0)) = V_t(\omega^*) \frac{\Pr[\omega = \omega^*]}{\Pr[\omega \geq \omega^*]} + V_t(m^*(\omega^*, 1)) \frac{\Pr[\omega > \omega^*]}{\Pr[\omega \geq \omega^*]}$$

$$\iff V_t(m^*(\omega^*, 0)) = V_t(\omega^*) \frac{m^*(\omega^*, 1) - m^*(\omega^*, 0)}{m^*(\omega^*, 1) - \omega^*} + V_t(m^*(\omega^*, 1)) \frac{m^*(\omega^*, 0) - \omega^*}{m^*(\omega^*, 1) - \omega^*}.$$

Analogously, the condition $\frac{\mathrm{d}V_t^*(\omega^*,0)}{\mathrm{d}t} \leq \frac{\mathrm{d}V_t^*(\omega^*,1)}{\mathrm{d}t}$ can be written as

$$V_t'(m^*(\omega^*, 0)) \geq V_t'(\omega^*) \frac{m^*(\omega^*, 1) - m^*(\omega^*, 0)}{m^*(\omega^*, 1) - \omega^*} + V_t'(m^*(\omega^*, 1)) \frac{m^*(\omega^*, 0) - \omega^*}{m^*(\omega^*, 1) - \omega^*}.$$

Thus, (15) states that if $V_t(m^*(\omega^*, 0))$ is equal to the convex combination of $V_t(\omega^*)$ and $V_t(m^*(\omega^*, 1))$, then the derivative $V_t'(m^*(\omega^*, 0))$ is greater than the convex combination of the derivatives $V_t'(\omega^*)$ and $V_t'(m^*(\omega^*, 1))$. When V is S-shaped, it is easy to see that this property is satisfied. □

Proof of Lemma 2. Observe that a signal induced by any set $X \subset C$ of permitted media outlets induces a monotone partition of $[0, 1]$, and thus $\mathcal{H}_C \subset \mathcal{H}_M$. Indeed, let $\underline{c}_X(\omega)$ and $\overline{c}_X(\omega)$ be the media outlets in X that are the closest to ω from above and from below:

$$\underline{c}_X(\omega) = \sup \Big(\{c \in X : c \leq \omega\} \cup \{0\} \Big) \quad \text{and} \quad \overline{c}_X(\omega) = \inf \Big(\{c \in X : c > \omega\} \cup \{1\} \Big).$$

By observing the media outlets in X, each citizen is informed that $\omega \in [\underline{c}_X(\omega), \overline{c}_X(\omega))$.

Conversely, let the set C of media outlets be $c_0 < \cdots < c_n$. Without loss of generality, assume that C contains the two uninformative media outlets $c_0 = 0$ and $c_n = 1$, which (almost) always endorse and criticize action $a = 1$, respectively. Then the distribution F of the state ω is a discrete distribution whose support is $\Omega = \{\omega_0, \ldots, \omega_{n-1}\}$ where $\omega_i = \mathbb{E}[\theta | \theta \in [c_i, c_{i+1})]$ for $i \in \{0, \ldots, n-1\}$. Consider a monotone partitional signal represented by a nondecreasing function $\xi : [0, 1] \to \mathbb{R}$. For each realization m of this signal that occurs with a strictly positive probability, define

$$\underline{x}(m) = \min \Big\{ \omega \in \Omega : \xi(\omega) = m \Big\} \quad \text{and} \quad \overline{x}(m) = \max \Big\{ \omega \in \Omega : \xi(\omega) = m \Big\}.$$

This signal induces a monotone partition of the finite set Ω. This partition of Ω can be implemented by a censorship policy that prohibits a media outlet $c \in C$ if and only if $c \in (\underline{x}(m), \overline{x}(m))$ for some realization m. □

References

Alonso, R., and O. Câmara (2016a): "Persuading Voters," *American Economic Review*, 106, 3590–3605.

——— (2016b): "Political Disagreement and Information in Elections," *Games and Economic Behavior*, 100, 390–412.

Bagnoli, M., and T. Bergstrom (2005): "Log-Concave Probability and Its Applications," *Economic Theory*, 26, 445–469.

Baron, D. P. (2006): "Persistent Media Bias," *Journal of Public Economics*, 90(1), 1–36.

Besley, T., and A. Prat (2006): "Handcuffs for the Grabbing Hand? Media Capture and Government Accountability," *American Economic Review*, 96(3), 720–736.

Blackwell, D. (1953): "Equivalent Comparisons of Experiments," *Annals of Mathematical Statistics*, 24, 265–272.

Chan, J., and W. Suen (2008): "A Spatial Theory of News Consumption and Electoral Competition," *Review of Economic Studies*, 75, 699–728.

Chiang, C.-F., and B. Knight (2011): "Media Bias and Influence: Evidence from Newspaper Endorsements," *Review of Economic Studies*, 78, 795–820.

Duffie, D., P. Dworczak, and H. Zhu (2017): "Benchmarks in Search Markets," *Journal of Finance*, 72, 1983–2044.

Dworczak, P. (2017): "Mechanism Design with Aftermarkets: Cutoff Mechanisms," mimeo.

Dworczak, P., and G. Martini (2019): "The Simple Economics of Optimal Persuasion," *Journal of Political Economy*, 127, 993–2048.

Edmond, C. (2013): "Information Manipulation, Coordination, and Regime Change," *Review of Economic Studies*, 80, 1422–1458.

Egorov, G., S. Guriev, and K. Sonin (2009): "Why Resource-poor Dictators Allow Freer Media: A Theory and Evidence from Panel Data," *American Political Science Review*, 103, 645–668.

Gehlbach, S., and K. Sonin (2014): "Government Control of the Media," *Journal of Public Economics*, 118, 163–171.

Gentzkow, M., and E. Kamenica (2016): "A Rothschild-Stiglitz Approach to Bayesian Persuasion," *American Economic Review, Papers & Proceedings*, 106, 597–601.

Gentzkow, M., and J. Shapiro (2006): "Media Bias and Reputation," *Journal of Political Economy*, 114(2), 280–316.

——— (2010): "What Drives Media Slant? Evidence from US Daily Newspapers," *Econometrica*, 78(1), 35–71.

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

Goldstein, I., and Y. Leitner (2018): "Stress Tests and Information Disclosure," *Journal of Economic Theory*, 177, 34–69.

Guo, Y., and E. Shmaya (2019): "The Interval Structure of Optimal Disclosure," *Econometrica*, 87, 653–675.

Johnson, J. P., and D. P. Myatt (2006): "On the Simple Economics of Advertising, Marketing, and Product Design," *American Economic Review*, 96, 756–784.

Kamenica, E., and M. Gentzkow (2011): "Bayesian Persuasion," *American Economic Review*, 101, 2590–2615.

Kolotilin, A. (2015): "Experimental Design to Persuade," *Games and Economic Behavior*, 90, 215–226.

——— (2018): "Optimal Information Disclosure: A Linear Programming Approach," *Theoretical Economics*, 13, 607–636.

Kolotilin, A., and H. Li (2019): "Relational Communication," mimeo.

Kolotilin, A., M. Li, T. Mylovanov, and A. Zapechelnyuk (2015): "Persuasion of a Privately Informed Receiver," mimeo.

Kolotilin, A., T. Mylovanov, A. Zapechelnyuk, and M. Li (2017): "Persuasion of a Privately Informed Receiver," *Econometrica*, 85, 1949–1964.

Kolotilin, A., and A. Wolitzky (2019): "Assortative Information Disclosure," mimeo.

Lewis, T. R., and D. Sappington (1994): "Supplying Information to Facilitate Price Discrimination," *International Economic Review*, 35, 309–327.

Lorentzen, P. (2014): "China's Strategic Censorship," *American Journal of Political Science*, 58, 402–414.

Mullainathan, S., and A. Shleifer (2005): "The Market for News," *American Economic Review*, 95(4), 1031–1053.

Orlov, D., P. Zryumov, and A. Skrzypach (2018): "Design of Macro-prudential Stress Tests," mimeo.

Ostrovsky, M., and M. Schwarz (2010): "Information Disclosure and Unraveling in Matching Markets," *American Economic Journal: Microeconomics*, 2, 34–63.

Prat, A., and D. Strömberg (2013): "The Political Economy of Mass Media," *Advances in Economics and Econometrics*, 2, 135–187.

Quah, J., and B. Strulovici (2012): "Aggregating the Single Crossing Property," *Econometrica*, 80, 2333–2348.

Rayo, L., and I. Segal (2010): "Optimal Information Disclosure," *Journal of Political Economy*, 118, 949–987.

Romanyuk, G., and A. Smolin (2019): "Cream Skimming and Information Design in Matching Markets," *American Economic Journal: Microeconomics*, 11, 250–276.

Suen, W. (2004): "The Self-Perpetuation of Biased Beliefs," *Economic Journal*, 114, 377–396.

Yamashita, T. (2018): "Optimal Public Information Disclosure by Mechanism Designer," mimeo.

Zapechelnyuk, A. (2019): "Optimal Quality Certification," *American Economic Review: Insights*, forthcoming.