Theoretical Economics 17 (2022), 561–585 1555-7561/20220561

Censorship as optimal persuasion

Anton Kolotilin
School of Economics, UNSW Sydney

Timofiy Mylovanov
Department of Economics, University of Pittsburgh

Andriy Zapechelnyuk
School of Economics and Finance, University of St Andrews

We consider a Bayesian persuasion problem where a sender's utility depends only on the expected state. We show that upper censorship that pools the states above a cutoff and reveals the states below the cutoff is optimal for all prior distributions of the state if and only if the sender's marginal utility is quasi-concave. Moreover, we show that it is optimal to reveal less information if the sender becomes more risk averse or the sender's utility shifts to the left. Finally, we apply our results to the problem of media censorship by a government.

Keywords. Bayesian persuasion, information design, censorship, media.

JEL classification. D82, D83, L82.

1. Introduction

Following Rayo and Segal (2010) and Kamenica and Gentzkow (2011), there has been a rapid growth of the literature on Bayesian persuasion (see Bergemann and Morris (2019) and Kamenica (2019) for excellent reviews). In a standard persuasion problem, a sender designs a signal about an uncertain state of the world to persuade a receiver. Much of

---

Anton Kolotilin: akolotilin@gmail.com
Timofiy Mylovanov: mylovanov@gmail.com
Andriy Zapechelnyuk: az48@st-andrews.ac.uk
We are grateful for discussions with Ming Li with whom we worked on the companion paper Kolotilin, Mylovanov, Zapechelnyuk, and Li (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, Li, Mylovanov, and Zapechelnyuk (2015). We thank the anonymous referees, Ricardo Alonso, Dirk Bergemann, Patrick Bolton, Alessandro Bonatti, Steven Callander, Odilon Câmara, Rahul Deb, Péter Esö, Florian Hoffman, Johannes Hörner, Roman Inderst, Emir Kamenica, Navin Kartik, Daniel Kräehmer, Hongyi Li, Marco Ottaviani, Mallesh Pai, Andrea Prat, Larry Samuelson, Ilya Segal, Joel Sobel, Konstantin Sonin, and Christopher Teh, as well as many conference and seminar participants, for helpful comments. We express a special gratitude to Ju Hu who made a substantial contribution by completing the proof of Theorem 1. Kolotilin 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. Mylovanov acknowledges financial support from the Office of Naval Research Multidisciplinary University Research Initiative Award N00014-17-1-2675 and from Kyiv School of Economics. Zapechelnyuk acknowledges financial support from the Economic and Social Research Council Grant ES/N01829X/1.

 https://doi.org/10.3982/TE4071

this literature focuses on a special *linear* case where the utilities depend only on the expected state. A wide range of applications 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 censor the states on one side of a cutoff and reveal the states on the other side of the cutoff. In this paper, we characterize necessary and sufficient conditions under which such censorship signals are optimal.

A linear persuasion problem is described by a prior distribution of a one-dimensional state and the sender's expected utility (under the receiver's optimal action) as a function of the expected state. We show that if the sender's utility is S-shaped (that is, the sender's marginal utility is quasi-concave) in the expected state, then, and only then, an *upper-censorship* signal that pools the states above a cutoff and reveals the states below the cutoff is optimal for all prior distributions of the state.[1] In addition, we perform comparative statics on the informativeness of the optimal censorship signal. When the sender's utility is S-shaped, the sender optimally reveals less information if she becomes more risk averse and if her utility shifts to the left.

To interpret an S-shaped utility, we consider a setting where a privately informed receiver chooses one of two actions: to accept or to reject a proposal. The sender reveals information about the proposal's value to the receiver. By accepting the proposal, the receiver forgoes a privately known outside option. If the sender wishes to maximize the probability that the receiver accepts the proposal, then the sender's utility is equal to the distribution function of the outside option. It is S-shaped (and thus upper censorship is optimal) if and only if the probability density of the outside option is unimodal. Furthermore, the sender is more risk averse (and thus optimally reveals less information) if and only if the density of the outside option decreases in the likelihood ratio order.

More generally, when the sender cares about a weighted sum of the receiver's utility and the probability that the proposal is accepted, the sender's utility is S-shaped for all possible weights if and only if the density of the outside option is log-concave. When this is the case, the sender optimally reveals less information if the weight she places on the receiver's utility falls and if each realization of the outside option decreases proportionally or by a constant, so that the receiver is more willing to accept the proposal.

We apply our results to the problem of media censorship. We consider a stylized setting with heterogeneous citizens and media outlets. Media outlets differ in their approval standards. A partially benevolent government influences citizens' actions by choosing which media outlets to permit and which ones to censor. Each permitted media outlet approves the government when the state is above its approval standard and criticizes it otherwise. In this context, upper censorship means that the government censors only media outlets with sufficiently high approval standards. We show that the government optimally censors more media outlets if the society experiences an ideology shock in favor of the government and if influencing the society's decisions becomes relatively more important than maximizing the citizens' individual welfare.

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

Related Literature. This paper is the first to provide necessary and sufficient conditions for the optimality of censorship that are prior-independent (as well as bias-independent in the setting with a privately informed receiver). This characterization unifies and generalizes sufficient conditions scattered across the literature. Under our conditions, the sender's problem is simply to find an optimal censorship cutoff. The independence of our conditions of the fine details of the environment enables a novel comparative statics analysis on the optimal censorship cutoff.

In related work, Kolotilin (2018) characterizes *prior-dependent* conditions for the optimality of interval disclosure, which includes censorship as a special case. Moreover, Alonso and Câmara (2016b) provide sufficient conditions for the optimality of censorship in a more specific setting with discrete states. They also perform comparative statics with respect to a parallel shift, which is a special case of a location–scale shift considered in our paper. Our comparative statics results with respect to the sender's risk aversion and bias are novel.

More broadly, this paper contributes to the literature on the linear persuasion problem highlighted by Kamenica and Gentzkow (2011).[2] Using Blackwell's theorem, Gentzkow and Kamenica (2016) and Kolotilin et al. (2017) characterize outcomes implementable by public signals and private persuasion mechanisms, respectively. By formulating the linear persuasion problem as a linear program, Kolotilin (2018), Dworczak and Martini (2019), and Dizdar and Kováč (2020) establish strong duality, thereby providing the verification tool for the optimality of a candidate signal. Finally, Arieli, Babichenko, Smorodinsky, and Yamashita (2021) and Kleiner, Moldovanu, and Strack (2021) characterize extreme solutions to a linear persuasion problem, thereby narrowing down the set of candidate optimal signals.

There is a diverse literature where censorship policies emerge as optimal signals in specific instances of the linear persuasion problem, starting from the prosecutor–judge example, as well as lobbying and product advertising examples, in Kamenica and Gentzkow (2011). Other contexts where censorship is optimal include grading policies (Ostrovsky and Schwarz (2010)), media control (Gehlbach and Sonin (2014), Ginzburg (2019), Gitmez and Molavi (2020)), 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 (2021)), online markets (Romanyuk and Smolin (2019)), attention management (Lipnowski, Mathevet, and Wei (2020), Bloedel and Segal (2021)), quality certification (Zapechelnyuk (2020)), and relational communication (Kolotilin and Li (2021)).

Finally, our application to media censorship contributes to the media economics literature, which we discuss in Section 5.3.

2. Model

A state of the world $\omega \in [0, 1]$ is a random variable whose prior probability distribution function F has a strictly positive probability density f on $[0, 1]$. A sender chooses a signal

[2]Nonlinear persuasion problems are studied, for example, by Rayo and Segal (2010), Goldstein and Leitner (2018), Guo and Shmaya (2019), and Kolotilin and Wolitzky (2020).

that reveals information about the state. A signal s is a random variable that is arbitrarily correlated with ω. For example, under *full disclosure*, s is perfectly correlated with ω, and under *no disclosure*, s is independent of ω.

The sender is an expected utility maximizer. The sender's Bernoulli utility V is a twice continuously differentiable function of the expected state $m = \mathbb{E}[\omega|s]$ induced by a signal s.[3] By Strassen's theorem, there exists a signal s that induces a probability distribution H of the expected state m if and only if the prior distribution F is a mean-preserving spread of H (see, for example, Kolotilin (2018) for details). We regard signals as identical if they induce the same distribution H of m. Thus, the sender's problem is to maximize the expected utility $\int_0^1 V(m)\,\mathrm{d}H(m)$ over distributions H such that F is a mean-preserving spread of H.

We assume that the sender's utility V is strictly increasing, unless stated otherwise. This is without loss of generality, because the solution to the sender's maximization problem is unaffected if we add a linear function to V. Indeed, $V(m) + bm$ is strictly increasing in m for $b > -\min_{z\in[0,1]} V'(z)$, where V' is a continuous derivative of V.

A signal is *upper censorship with cutoff* $\omega^* \in [0, 1]$ if it reveals the states below ω^* and pools the states above ω^*. The expected state of the *pool* $[\omega^*, 1]$ is denoted by $m^* = \mathbb{E}[\omega|\omega \geq \omega^*]$, which is a function of ω^*. In particular, full disclosure and no disclosure are upper-censorship signals with cutoffs $\omega^* = 1$ and $\omega^* = 0$. Under upper censorship with cutoff ω^*, the sender's expected utility is

$$W(\omega^*) = \int_0^{\omega^*} V(\omega)f(\omega)\,\mathrm{d}\omega + \int_{\omega^*}^1 V(m^*)f(\omega)\,\mathrm{d}\omega. \tag{1}$$

A *lower-censorship signal* is defined symmetrically: it pools the states below a specified cutoff and reveals the states above this cutoff. As the results for lower censorship are analogous, we focus on upper censorship throughout this paper.

3. Results

A function V is said to be *(strictly) S-shaped* on $[0, 1]$ if there exists $x \in [0, 1]$ such that V is (strictly) convex on $[0, x]$ and (strictly) concave on $[x, 1]$, or, equivalently, if its derivative V' is (strictly) quasi-concave on $[0, 1]$.

We say that upper censorship is *(uniquely) optimal* if there exists a (unique) optimal signal that is an upper censorship with some cutoff $\omega^* \in [0, 1]$. When upper censorship with cutoff ω^* is (uniquely) optimal, the cutoff ω^* must be a (unique) point that maximizes the function W given by (1).

We now relate the optimality of upper censorship to the S-shapedness of V.

Theorem 1. *If V is (strictly) S-shaped on* $[0, 1]$*, then upper censorship is (uniquely) optimal for each density f. Conversely, if V is not (strictly) S-shaped on* $[0, 1]$*, then there exists a density f such that upper censorship is not (uniquely) optimal.*

[3] Section 4.1 provides a micro-founded model with a receiver where the sender's Bernoulli utility depends only on the expected state. More generally, this assumption holds in sender–receiver models where the receiver's optimal action depends only on the expected state and the sender's utility is linear in the state.

To provide the intuition for Theorem 1, we first discuss how the optimal censorship cutoff is determined, and then show that the upper-censorship signal with this cutoff is optimal among all signals. The optimal cutoff ω^* maximizes the sender's expected utility W. The derivative of W is

$$W'(\omega^*) = \big(V(\omega^*) - V(m^*)\big) f(\omega^*) + V'(m^*) \frac{\mathrm{d}m^*}{\mathrm{d}\omega^*}\big(1 - F(\omega^*)\big). \tag{2}$$

Intuitively, when increasing the cutoff by $\mathrm{d}\omega^*$, the sender experiences two effects. First, the sender's utility decreases by $V(m^*) - V(\omega^*)$ for the mass $f(\omega^*)\,\mathrm{d}\omega^*$ of the marginal states that are now separated from the pool. Second, when the lowest states $[\omega^*, \omega^* + \mathrm{d}\omega^*)$ of the pool are separated, the expected state m^* of the pool increases by $\mathrm{d}m^*$. This in turn increases the sender's utility by $V'(m^*)\,\mathrm{d}m^*$ for the mass $1 - F(\omega^*)$ of the pooled states. When the sender's optimal cutoff is interior, the first-order condition equates these two effects, so $W'(\omega^*) = 0$. In turn, the second-order condition requires that W' crosses the horizontal axis from above at ω^*. When V is S-shaped, W' is indeed single-crossing from above, so an optimal cutoff solves a standard quasi-concave maximization problem.

Lemma 1. *If V is (strictly) S-shaped on $[0, 1]$, then W is (strictly) quasi-concave for each density f. Conversely, if V is not (strictly) S-shaped on $[0, 1]$, then there exists a density f such that W is not (strictly) quasi-concave.*

To gain the intuition for Lemma 1, we use

$$\frac{\mathrm{d}m^*}{\mathrm{d}\omega^*} = \frac{\mathrm{d}}{\mathrm{d}\omega^*}\left(\frac{1}{1 - F(\omega^*)}\int_{\omega^*}^{1} z f(z)\,\mathrm{d}z\right) = \frac{(m^* - \omega^*) f(\omega^*)}{1 - F(\omega^*)}$$

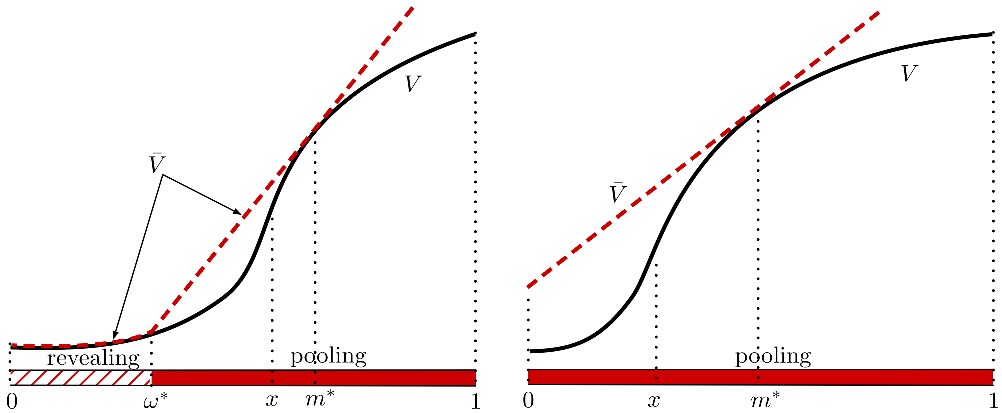
to rewrite (2) as

$$W'(\omega^*) = f(\omega^*)\big[V(\omega^*) - \big(V(m^*) + V'(m^*)(\omega^* - m^*)\big)\big]. \tag{3}$$

Thus, $W'(\omega^*) \leq 0$ holds if and only if the difference at ω^* between the tangent line to V at m^* and the utility V itself is nonnegative. Figure 1(a) and (b) shows the cases where $W'(\omega^*) = 0$ with $\omega^* \in (0, 1)$ and $W'(\omega^*) < 0$ with $\omega^* = 0$. Clearly, increasing ω^* increases m^*, so the tangent line at m^* becomes flatter and thus remains above V at ω^*, ensuring that $W'(\omega^*) \leq 0$ still holds. This shows that W is quasi-concave.

Inspecting the first-order condition, $W'(\omega^*) = 0$ in Figure 1(a) and $W'(\omega^*) < 0$ in Figure 1(b), we can see that the optimal cutoff ω^* and the expected state m^* of the pool must lie on the opposite sides of an inflection point x of V, when V is S-shaped.

Lemma 2. *Let V be S-shaped on $[0, 1]$ and let $\omega^* \in [0, 1]$ (respectively, $\omega^* \in (0, 1)$) be a maximum of W. There exists $x \in [\omega^*, m^*]$ (respectively, $x \in (\omega^*, m^*)$) such that V is convex on $[0, x]$ and concave on $[x, 1]$.*



(a) The case of $\omega^* \in (0, 1)$ (b) The case of $\omega^* = 0$

FIGURE 1. Upper-censorship signals.

We can now outline the intuition for Theorem 1. To understand why upper censorship is optimal among all signals, consider again an S-shaped V and the optimal censorship cutoff ω^*, as shown in Figure 1. Let $\bar{V}(m)$ be given by $V(m)$ for $m < \omega^*$ and $V(m^*) + V'(m^*)(m - m^*)$ for $m \geq \omega^*$, as illustrated by the dashed curve in Figure 1. We interpret $\bar{V}$ as the Bernoulli utility of a third party. Since $\bar{V}$ is convex, the party is risk loving and thus her optimal signal is full disclosure, which induces the riskiest lottery over m. A moment's reflection shows that the party's value of her optimal signal is $W(\omega^*)$, which coincides with the sender's value of upper censorship with cutoff ω^*. As the sender's utility V is not higher than the party's utility $\bar{V}$, this upper-censorship signal must, therefore, be optimal for the sender.

To build the intuition for the converse of Theorem 1, suppose, for simplicity, that V has a finite number of inflection points. If V is not S-shaped, then there is an interval of states where V' is quasi-convex and not quasi-concave. Thus, it is possible to choose a density f such that it is strictly positive only on this interval, and the unique optimal signal is lower censorship with an interior cutoff, so upper censorship is suboptimal. Since V is continuous, upper censorship remains suboptimal for any strictly positive density sufficiently close to f.

It is straightforward to specialize Theorem 1 to extreme forms of upper censorship.

COROLLARY 1. *If V is (strictly) convex/concave on* $[0, 1]$*, then full/no disclosure is (uniquely) optimal for each density f. Conversely, if V is not (strictly) convex/concave on* $[0, 1]$*, then there exists a density f such that full/no disclosure is not (uniquely) optimal.*

Theorem 1 enables a comparative statics analysis on Blackwell informativeness of the optimal signal. When V is S-shaped, the sender solves a quasi-concave maximization problem, so the set of optimal censorship cutoffs is an interval. When comparing two upper-censorship signals, we say that the sender reveals less information if both bounds of the interval of optimal cutoffs are smaller. In particular, in the case of *strictly*

S-shaped V, the optimal cutoff is unique, so the sender reveals less information if and only if a smaller interval of states is revealed.

First, we show that the sender reveals less information if she is more risk averse in the Arrow–Pratt sense. Let V_1 and V_2 be S-shaped. We say that V_2 is *more risk averse* than V_1 if V_2 is a concave transformation of V_1 or, equivalently, if for all $z \in [0, 1]$,

$$-\frac{V_2''(z)}{V_2'(z)} \geq -\frac{V_1''(z)}{V_1'(z)}. \tag{4}$$

Proposition 1. *Let V be S-shaped on* $[0, 1]$. *The sender optimally reveals less information if V becomes more risk averse.*

Second, we show that the sender reveals less information if her Bernoulli utility shifts to the left in the location–scale sense.[4] For this analysis, we define V on the real line $\mathbb{R}$, rather than on the unit interval (as we have implicitly done so far). Accordingly, V is said to be S-shaped on $\mathbb{R}$ if V' is quasi-concave on $\mathbb{R}$.

We say that V_2 is a (location–scale) *shift to the left* of V_1 if, for all $z \in \mathbb{R}$,

$$V_2(z) = V_1\left(\frac{z-\alpha}{\beta}\right), \quad \text{where } \alpha \leq 0 \text{ and } 0 < \beta \leq 1. \tag{5}$$

That is, the graph of V_2 is obtained from the graph of V_1 by shifting each point to the left from $(z, V_1(z))$ to $(\alpha + \beta z, V_1(z))$, where $\alpha + \beta z \leq z$ for all $z \in [0, 1]$, because $\alpha \leq 0$ and $\beta \leq 1$. Moreover, if V_1 is S-shaped, then so is V_2, because $\beta > 0$.

Proposition 2. *Let V be S-shaped on $\mathbb{R}$. The sender optimally reveals less information if V shifts to the left.*

Figure 2 illustrates how the optimal censorship cutoff decreases when the sender becomes more risk averse and her utility shifts to the left.[5] We defer the economic intuition to Section 4.2.

The comparative statics results hold in the strong form with similar proofs. Specifically, suppose that V is strictly S-shaped and the optimal censorship cutoff ω^* is interior. Then the sender reveals strictly less information in that ω^* strictly decreases (i) if V becomes strictly more risk averse in that the inequality in (4) is strict, or (ii) if V strictly shifts to the left in that $(\alpha, \beta) \neq (0, 1)$ in (5).

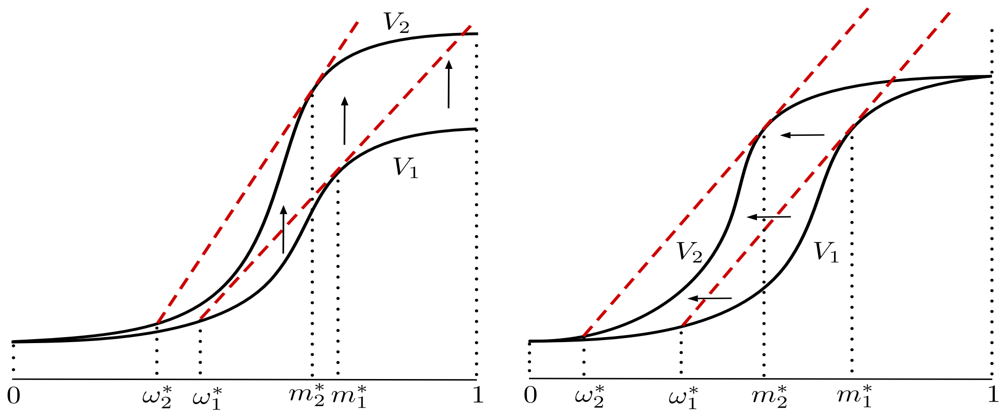
4. Additive representation

This section focuses on the *additive representation* of V given by

$$V(m) = G(m) + \rho \int_0^m G(r)\,\mathrm{d}r, \tag{6}$$

[4] Clearly, a vertical linear transformation of V defined as $\alpha + \beta V$, with $\beta > 0$, does not change the solution to the sender's maximization problem, thus having no effect on the optimal signal.

[5] In general, Proposition 1 neither implies nor is implied by Proposition 2. But in the special case of increasing risk aversion in that $-V''(z)/V'(z)$ is increasing in z, the sender becomes more risk averse if her utility shifts to the left, so Proposition 2 follows from Proposition 1.



(a) V_2 is more risk averse than V_1 (b) V_2 is a shift to the left of V_1

FIGURE 2. Comparative statics.

where $\rho \in \mathbb{R}$ is a parameter, and G is a probability distribution function that has a strictly positive and continuously differentiable density g on $[0, 1]$. To motivate this representation, Section 4.1 presents a micro-founded model with a privately informed receiver who chooses between two actions. Section 4.2 sets the stage by reinterpreting the results of Section 3 in the simple case where $\rho = 0$. Section 4.3 specifies explicit conditions on ρ and G for upper censorship to be optimal and analyzes how changes in ρ and G affect Blackwell informativeness of the optimal signal.[6]

4.1 *Microeconomic foundation*

There are two players: a sender and a receiver. The receiver chooses whether to accept a proposal ($a = 1$) or to reject it ($a = 0$). The proposal has an uncertain value (state) $\omega \in [0, 1]$ for the receiver. By accepting the proposal, the receiver forgoes an outside option worth $r \in [0, 1]$, which is the receiver's private information. The state and outside option are independent random variables whose probability distributions F and G have strictly positive probability densities f and g, with g being continuously differentiable.

The receiver's and sender's utilities from $a = 0$ are normalized to zero, and their utilities from $a = 1$ are given by

$$u(\omega, r) = \omega - r \quad \text{and} \quad v(\omega, r) = 1 + \rho(\omega - r), \tag{7}$$

where $\rho \in \mathbb{R}$ is an *alignment* parameter. That is, the sender's utility is a weighted sum of the action $a = 1$ and the receiver's utility. The alignment parameter ρ captures the relative weight that the sender assigns to the receiver's utility. A sender with such preferences may correspond to a partially benevolent government as in Section 5 or to a recommender system that cares about both producer and consumer welfare.

[6]When $\rho \neq 0$, V does not have to be S-shaped if G is S-shaped, and V does not have to shift to the left if G does.

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

Conditional on the expected state m, the receiver's expected utility from $a = 1$ is $m - r$. So the receiver optimally chooses $a = 1$ if and only if $r \leq m$. Consequently, conditional on m, the sender's posterior expected utility is

$$V(m) = \int_0^m \big(1 + \rho(m - r)\big) g(r)\, \mathrm{d}r = G(m) + \rho \int_0^m G(r)\, \mathrm{d}r.$$

That is, the sender's Bernoulli utility function is indeed given by (6).

4.2 *State-independent case*

A special case that appears in much of the persuasion literature is where the sender wishes to maximize the probability that the receiver accepts the proposal. This is the case of $\rho = 0$, so $v(\omega, r) = 1$ and $V(m) = G(m)$. Thus, V is S-shaped if and only if the density g of the outside option is unimodal, where the inflection point of V is the mode of g. By Theorem 1, this unimodality is necessary and sufficient for the optimality of upper censorship.

In this case, the comparative statics results of Propositions 1 and 2 are also easy to interpret. Let g_1 and g_2 be unimodal densities of distributions G_1 and G_2 of the outside option. First, observe that G_2 is more risk averse than G_1 if and only if g_2/g_1 is decreasing, meaning that the distribution G_2 of the outside option is smaller in the likelihood ratio order than the distribution G_1. Second, observe that G_2 is a (location–scale) shift to the left of G_1 if and only if each r is replaced with a smaller value $\alpha + \beta r$, meaning that each realization of the outside option is smaller proportionally and by a constant. Thus, the sender optimally reveals less information if the receiver becomes more willing to accept the proposal, in that the distribution of the outside option decreases either in the likelihood-ratio order (Proposition 1) or in the location–scale shift order (Proposition 2).

To build intuition, consider the limit case where the density g is symmetric and highly concentrated around its mode $x > \mathbb{E}[\omega]$, meaning that the outside option is close to x with high probability. In this case the optimal censorship cutoff $\omega^* < x$ is such that the expected state of the pool $m^* = \mathbb{E}[\omega|\omega \geq \omega^*]$ is slightly above x, so that the receiver is very likely to accept the proposal. On the one hand, a further increase in m^* reduces the probability that the state belongs to the pooling interval $[\omega^*, 1]$, but only negligibly increases the acceptance probability conditional on $\omega \in [\omega^*, 1]$. On the other hand, setting $m^* \leq x$ drops the acceptance probability to below a half. This illustrates Lemma 2 asserting that $x \in (\omega^*, m^*)$.

Now suppose that the distribution of the outside option either decreases in the likelihood-ratio order or in the location–scale shift order, so that the outside option is now close to some $\tilde{x} < x$ with high probability. By the intuition provided above, the optimal censorship cutoff should then decrease so that the expected state of the pool falls from slightly above x to slightly above $\tilde{x}$. This illustrates Propositions 1 and 2.

4.3 *Weighted case*

We now provide necessary and sufficient conditions for the optimality of upper censorship for the additive representation of V given by (6), without restrictions on ρ.

A probability density g is *(strictly) log-concave* on a given interval if $\ln g$ is (strictly) concave on that interval. Log-concavity is a common assumption in a variety of economic applications, such as voting, signalling, and monopoly pricing (see Section 7 in Bagnoli and Bergstrom (2005)). Log-concave densities exhibit nice properties, such as unimodality and hazard rate monotonicity. Many familiar probability density functions are log-concave (see Table 1 in Bagnoli and Bergstrom (2005)).

The next theorem connects the optimality of upper censorship and the log-concavity of the density g.

Theorem 2. *Let V be given by (6). If g is (strictly) log-concave on* $[0, 1]$*, then upper censorship is (uniquely) optimal for each f and each ρ. Conversely, if g is not (strictly) log-concave on* $[0, 1]$*, then there exist f and ρ such that upper censorship is not (uniquely) optimal.*

The proof of Theorem 2 is immediate by Theorem 1 and the following lemma that links the S-shapedness of V and the log-concavity of g.[7]

Lemma 3. *Let V be given by (6). If g is (strictly) log-concave on* $[0, 1]$*, then V is (strictly) S-shaped on* $[0, 1]$*. Conversely, if g is not (strictly) log-concave on* $[0, 1]$*, then there exists ρ such that V is not (strictly) S-shaped on* $[0, 1]$.

Proof. By (6), using the assumption that g is strictly positive on $[0, 1]$, we have, for $m \in [0, 1]$,

$$V''(m) = g'(m) + \rho g(m) = g(m)\left(\frac{g'(m)}{g(m)} + \rho\right). \tag{8}$$

If g is (strictly) log-concave on $[0, 1]$, then g'/g is (strictly) decreasing on $[0, 1]$. So V'' (strictly) crosses the horizontal axis from above at most once on $[0, 1]$ and, thus, V is (strictly) S-shaped on $[0, 1]$. Conversely, if g is not (strictly) log-concave on $[0, 1]$, then there exist $m_1 < m_2$ such that g'/g is strictly (weakly) increasing on $[m_1, m_2]$. Choosing $\rho = -g'(m)/g(m)$ for some $m \in (m_1, m_2)$, we obtain that V is not (strictly) S-shaped on $[0, 1]$. □

The conditions for the optimality of lower censorship are symmetric to those presented in Theorem 2. Specifically, if V is given by (6) and g is log-convex, then lower censorship is optimal. Thus, if g is both log-concave and log-convex (so it is exponential), then an optimal signal is both upper censorship and lower censorship. The only two signals with this property are full disclosure and no disclosure. This shows that the

[7]Lemma 3 and Theorem 2 extend to a more general case of $v(\omega, r) = \nu(r)(1 + \rho(r)(\omega - r))$, where ρ is continuous, and ν is strictly positive and continuously differentiable. In that case, V is (strictly) S-shaped if ν and g are (strictly) log-concave and ρ is decreasing.

optimal signal is polarized between full disclosure and no disclosure (as, for example, in Lewis and Sappington (1994) and Johnson and Myatt (2006)) only in the knife edge case where the receiver's outside option has an exponential distribution.

Corollary 2. *Let V be given by (6). If g is exponential, in that $g'(r)/g(r) = -\lambda$ for some $\lambda \in \mathbb{R}$, then, for each density f, full disclosure is uniquely optimal for $\rho > \lambda$, no disclosure is uniquely optimal for $\rho < \lambda$, and all signals are optimal for $\rho = \lambda$. Conversely, if g is not exponential, then there exist f and ρ such that neither full disclosure nor no disclosure is optimal.*

We now adapt Propositions 1 and 2 to obtain comparative statics results for the additive representation.

Proposition 3. *Let V be given by (6) and let g be log-concave. The sender optimally reveals less information*

(i) if the alignment parameter ρ decreases

(ii) if the distribution G of the outside option shifts to the left.

The proof of part (i) shows that if ρ decreases, then V becomes more risk averse. Intuitively, if the sender puts a lower weight ρ on the receiver's preferences, she optimally endows the receiver with a lower utility by disclosing less information.

The proof of part (ii) shows that if G shifts to the left, then V also shifts to the left and, in addition, becomes more risk averse. The intuition for part (ii) is the same as in the state-independent case of Section 4.2.

5. Application to media censorship

In this section, we present a model of media censorship by the government and show that it can be represented as a persuasion problem with a privately informed receiver in Section 4.1. 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.

5.1 *Model*

There is a continuum of heterogeneous citizens indexed by $r \in [0, 1]$ distributed with G that has a strictly positive and twice continuously differentiable density g. The state of the world $\omega \in [0, 1]$ has a distribution F that has a strictly positive density f. Each citizen chooses between $a = 0$ and $a = 1$. The utility of a citizen of type r is given by

$$u(a_r, \omega, r) = (\omega - r)a_r,$$

where $a_r \in \{0, 1\}$ is the citizen's action.

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. The state of the world ω is a common benefit of action $a = 1$, whereas r is a type-specific cost of this action. Thus, the state ω can be interpreted as the government's quality or valence, and the type r can be interpreted as a citizen's ideological position or preference parameter.

There is a government that cares about a weighted average of the social utility and the government's intrinsic benefit from the aggregate action. For a given state ω, the government's utility is given by

$$B(\bar{a}) + \rho \int_0^1 u(a_r, \omega, r) g(r) \, \mathrm{d}r,$$

where $\bar{a} = \int_0^1 a_r g(r) \, \mathrm{d}r$ is the aggregate action in the society and B is a twice continuously differentiable function with a strictly positive derivative b. Parameter ρ captures the alignment of the government's preferences with those of the citizens, and the term $B(\bar{a})$ captures the government's intrinsic benefit from the aggregate action as well as possible externalities of the citizens' actions. Since the marginal benefit b from increasing the aggregate action is strictly positive and, in addition, more citizens prefer action $a = 1$ when the state ω is higher, a high state is good news for the government.

Citizens obtain information about the unobservable state ω through media outlets. There is a continuum of media outlets indexed by $c \in [0, 1]$. Each media outlet c approves the pro-government action $a = 1$ if $\omega \geq c$ and criticizes it if $\omega < c$. We thus can interpret c as the media outlet's standard of approval. A media outlet with a higher approval standard c is more opposing to the government because it criticizes the pro-government action on a larger set of states.

The government can censor media outlets to stop them from broadcasting. Censored media outlets are uninformative to the public. The government's censorship policy is a set of the media outlets $X \subset [0, 1]$ that are censored. The other media outlets in $[0, 1] \backslash X$ continue to broadcast. The government's censorship of media outlets can take various forms, including banning access to internet sites, withdrawing licenses, and even arresting editors and journalists using broadly formulated legislation on combating extremism. Importantly, the government implements a censorship policy before the state is realized. This assumption reflects that daily news coverage is the privilege of the media, and the government cannot routinely interfere in this process.[8]

A government censorship policy is *upper censorship* with cutoff $\omega^* \in [0, 1]$ if it censors all sufficiently opposing media outlets. Specifically, media outlets whose approval standards are below ω^* are permitted and the rest are censored, so $X = [\omega^*, 1]$. Note that *full censorship* is an extreme form of upper censorship where all media outlets are censored ($\omega^* = 0$), while *free media* is a degenerate form of upper censorship where (almost) no media outlets are censored ($\omega^* = 1$).

The timing is as follows. First, the government chooses a set $X \subset [0, 1]$ of censored media outlets. Second, the state ω is realized, and each permitted media outlet approves or criticizes action $a = 1$ according to its approval standard. Finally, each citizen

[8]A similar assumption is made in Gehlbach and Sonin (2014) and Gitmez and Molavi (2020). If, instead, the government could choose a censorship policy after having observed the state, then the state would be fully revealed in equilibrium, due to the unraveling argument of Milgrom (1981).

observes messages from all permitted media outlets, updates his beliefs about ω, and chooses an action.

5.2 *Media censorship as persuasion*

We now show that the media censorship problem can be formulated as a persuasion problem, in which the government is a sender and a representative citizen is a receiver.

Observe that a government's upper-censorship policy with cutoff ω^* is equivalent to an upper-censorship signal that reveals the states below ω^* and pools the states above ω^*. Consequently, when upper censorship is optimal among all signals, it is also optimal among signals induced by censoring media outlets.

Let m be the expected state induced by messages from the permitted media outlets. A citizen of type r chooses $a_r = 1$ if and only if $r \leq m$. Consequently, the aggregate action $\bar{a}$ is the mass of all citizens whose types are at most m, so $\bar{a} = G(m)$.

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

$$\begin{aligned} V(m) &= \mathbb{E}\left[B(\bar{a}) + \rho \int_0^1 u(a_r, \omega, r) g(r)\, \mathrm{d}r \,\middle|\, m\right] \\ &= B\big(G(m)\big) + \rho \int_0^m (m-r) g(r)\, \mathrm{d}r = \int_0^{G(m)} b(r)\, \mathrm{d}r + \rho \int_0^m G(r)\, \mathrm{d}r. \end{aligned} \tag{9}$$

We thus obtain the persuasion problem in Section 2, where the sender's Bernoulli utility is given by (9). By Theorem 1, upper censorship is optimal among all censorship policies if V is S-shaped.

Consider two special cases of interest: *majoritarian* and *proportional*. In the majoritarian case, the government cares only about the aggregate action, so $\rho = 0$ and $V(m) = B(G(m))$. For example, this can reflect the government's utility when it is supported by the fraction $G(m)$ of voters in an upcoming election.

THEOREM 3. *Let V be given by (9) with $\rho = 0$. If b and g are (strictly) log-concave on $[0, 1]$, then upper censorship is (uniquely) optimal.*

Theorem 3 shows that an upper-censorship policy is optimal if both the population density g and the marginal benefit b from convincing citizens to choose $a = 1$ are log-concave and thus unimodal. This condition on b holds if the government's top priority is to reach a certain approval threshold such as a simple majority, so that B is a smooth approximation of a step function.

The second special case of interest is where the government's benefit is proportional to the share of supporters, $B(\bar{a}) = \bar{a}$. Then V is the same as (6) in Section 4 and, thus, upper censorship is optimal when g is log-concave, by Theorem 2.

Propositions 2 and 3 imply that in both the majoritarian and proportional cases, the government optimally censors more media outlets if the citizens are more supportive of the government, so the cutoff approval standard ω^* decreases if the distribution of the citizens' types is shifted to the left. Moreover, in the proportional case, the government censors more media outlets if it cares more about the approval and less about the social welfare, so ω^* decreases if ρ becomes smaller.

5.3 *Discussion*

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 a plurality of media outlets and on transaction costs of bribing the media. In related work, Gehlbach and Sonin (2014) consider a setting with a government-influenced monopoly media outlet that trades off mobilizing the population for some collective goal and collecting the revenue from subscribers who demand informative news.[9] 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 heterogeneous media outlets. As a result, the government optimally discriminates media outlets by permitting sufficiently supportive ones and banning the remaining ones.

As in Suen (2004), Chan and Suen (2008), and Chiang and Knight (2011), we assume that media outlets use binary approval policies that communicate only whether the state is above some standard. This assumption reflects a cursory reader's preference for simple messages such as positive or negative opinions and yes or no recommendations. Our results, however, do not rely on this assumption. As we have shown, when the government's utility is S-shaped in the expected state, as in the majoritarian and proportional cases, upper-censorship policies are optimal among all possible signals about the state of the world. Thus, the government cannot benefit from introducing new media outlets or using more complex forms of media control such as aggregating information from multiple media outlets and possibly adding noise.

In our model, each citizen observes messages from all permitted media outlets. However, among all these media outlets, only one is pivotal for a citizen's choice between two actions. Therefore, our results are not affected if each citizen could follow only his most preferred media outlet, as in Chan and Suen (2008). More generally, the government cannot benefit from information discrimination of citizens, as follows from Kolotilin et al. (2017).

Our results can be easily extended to the case of a finite number of media outlets. When the government's utility is S-shaped in the expected state, it is again optimal to censor all sufficiently opposing media outlets and permit the rest.

In our model, each citizen's optimal action does not depend on what other citizens do. Otherwise the citizens' optimal actions would constitute an equilibrium of the game induced by the available information. Imposing some equilibrium selection criterion,

[9] 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).

we can still define the government's utility as a function of the expected state and apply our results when the citizens' and government's utilities are linear in the state.

There are other aspects that can be relevant in media economics. Citizens can incur some cost of following media outlets. 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. Moreover, it can be costly for the government to censor media outlets. So another important question is how the government should prioritize censoring. These extensions are nontrivial and left for future research.

Appendix

We first prove a few lemmas, which are used in Theorem 1 and other results.

Lemma 4. *Let* $\omega^* \in [0, 1]$ *and* $m^* = \mathbb{E}[\omega|\omega \geq \omega^*]$. *Then*

$$W'(\omega^*) = f(\omega^*) \int_{\omega^*}^{m^*} V''(z)(z - \omega^*)\,\mathrm{d}z.$$

Proof. The proof follows from (3) and integration by parts:

$$\begin{aligned}\int_{\omega^*}^{m^*} V''(z)(z - \omega^*)\,\mathrm{d}z &= V'(z)(z - \omega^*)\Big|_{\omega^*}^{m^*} - \int_{\omega^*}^{m^*} V'(z)\,\mathrm{d}z \\ &= V'(m^*)(m^* - \omega^*) - \big(V(m^*) - V(\omega^*)\big).\end{aligned}$$

□

Lemma 5. *Let* $\omega_1 < \omega_2$ *and* $\omega_1 < m_1 \leq m_2$. *If there exists* $x \in [0, m_1]$ *such that* V *is convex on* $[0, x]$ *and concave on* $[x, 1]$, *then*

$$\int_{\omega_1}^{m_1} V''(z)(z - \omega_1)\,\mathrm{d}z \leq 0 \quad \Longrightarrow \quad \int_{\omega_2}^{m_2} V''(z)(z - \omega_2)\,\mathrm{d}z \leq 0. \tag{10}$$

If, in addition, $\omega_2 \neq m_2$ *and* V *is strictly concave on* $[m_1 - \varepsilon, 1]$ *for some* $\varepsilon > 0$, *then*

$$\int_{\omega_1}^{m_1} V''(z)(z - \omega_1)\,\mathrm{d}z \leq 0 \quad \Longrightarrow \quad \int_{\omega_2}^{m_2} V''(z)(z - \omega_2)\,\mathrm{d}z < 0. \tag{11}$$

Proof. For $\omega_2 \geq x$, the conclusion of (10) follows regardless of its hypothesis. For $\omega_2 < x$, (10) follows from

$$\begin{aligned}\int_{\omega_2}^{m_2} V''(z)(z - \omega_2)\,\mathrm{d}z &\leq \int_{\omega_2}^{m_1} V''(z)(z - \omega_2)\,\mathrm{d}z && (12) \\ &= \int_{\omega_2}^{m_1} V''(z)(z - \omega_1)\frac{z - \omega_2}{z - \omega_1}\,\mathrm{d}z \\ &\leq \frac{x - \omega_2}{x - \omega_1} \int_{\omega_2}^{m_1} V''(z)(z - \omega_1)\,\mathrm{d}z && (13)\end{aligned}$$

$$\leq \frac{x-\omega_2}{x-\omega_1}\int_{\omega_1}^{m_1} V''(z)(z-\omega_1)\,\mathrm{d}z \tag{14}$$

$$\leq 0, \tag{15}$$

where (12) holds because $x \leq m_1 \leq m_2$, (13) holds because $(z-\omega_2)/(z-\omega_1)$ is strictly increasing in z and $V''(z) \geq (\leq) 0$ for $z < (>) x$, (14) holds because $\omega_1 < \omega_2 < x$, and (15) holds by the hypothesis of (10).

For $\omega_2 \geq x$ with $\omega_2 \neq m_2$, the conclusion of (11) follows regardless of its hypothesis. For $\omega_2 < x$, (11) follows because the inequality in (13) is strict. □

Proof of Lemma 1. If V is S-shaped, then, by Lemma 4, $W'(\omega_1^*) < 0$ implies that $x < m_1^*$. Thus, W is quasi-concave, because $W'(\omega_1^*) < 0$ implies $W'(\omega_2^*) \leq 0$ for $\omega_1^* < \omega_2^*$, by (10) in Lemma 5.

Similarly, if V is strictly S-shaped, then, by Lemma 4, $W'(\omega_1^*) \leq 0$ for $\omega_1^* < m_1^*$ implies that $x < m_1^*$. Thus, W is strictly quasi-concave, because $W'(\omega_1^*) \leq 0$ implies $W'(\omega_2^*) < 0$ for $\omega_1^* < \omega_2^* < 1$, by (11) in Lemma 5.

Suppose that V is not S-shaped. Then there exist $\omega_1^* < m_1^* < \omega_2^* < m_2^*$ and a strictly positive density f such that $V''(z) < 0$ for $z \in [\omega_1^*, m_1^*]$, $V''(z) > 0$ for $z \in [\omega_2^*, m_2^*]$, $m_1^* = \mathbb{E}[\omega|\omega \geq \omega_1^*]$, and $m_2^* = \mathbb{E}[\omega|\omega \geq \omega_2^*]$. Then $W'(\omega_1^*) < 0 < W'(\omega_2^*)$ by Lemma 4, showing that W is not quasi-concave.

Finally, suppose that V is S-shaped but not strictly S-shaped. Then there exist $\omega_1^* < m_1^* < \omega_2^* < m_2^*$, and a strictly positive density f such that $V''(z) = 0$ for $z \in [\omega_1^*, m_2^*]$, $m_1^* = \mathbb{E}[\omega|\omega \geq \omega_1^*]$, and $m_2^* = \mathbb{E}[\omega|\omega \geq \omega_2^*]$. Then $W'(\omega^*) = 0$ for all $\omega^* \in [\omega_1^*, \omega_2^*]$ by Lemma 4, showing that W is not strictly quasi-concave. □

Proof of Lemma 2. By Lemma 1, W is quasi-concave, so $W'(\tilde{\omega}^*) \geq (\leq) 0$ for $\tilde{\omega}^* < (>) \omega^*$ or, equivalently, by Lemma 4,

$$\int_{\tilde{\omega}^*}^{\tilde{m}^*} V''(z)\big(z-\tilde{\omega}^*\big)\,\mathrm{d}z \geq (\leq) 0 \quad \text{for } \tilde{\omega}^* < (>) \omega^*, \tag{16}$$

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

First, suppose that $\omega^* \in (0,1)$. Then (16) implies that $\int_{\omega^*}^{m^*} V''(z)(z-\omega^*)\,\mathrm{d}z = 0$. Thus, since $\omega^* < m^*$ and V'' is single-crossing from above, there must exist $x \in (\omega^*, m^*)$ such that $V''(z) \geq (\leq) 0$ for $z < (>) x$.

Next suppose that $\omega^* = 0$. Then (16) implies that $\int_0^{m^*} V''(z) z\,\mathrm{d}z \leq 0$. Thus, since $\omega^* < m^*$ and V'' is single-crossing from above, there must exist $x \in [\omega^*, m^*)$ such that $V''(z) \geq (\leq) 0$ for $z < (>) x$.

Finally, suppose that $\omega^* = 1$. Then (16) implies that $\int_{\tilde{\omega}^*}^{\tilde{m}^*} V''(z)(z-\tilde{\omega}^*)\,\mathrm{d}z \geq 0$ for all $\tilde{\omega}^* < 1$. Since $\tilde{m}^*$ is a continuous increasing function of $\tilde{\omega}^*$ that satisfies $\tilde{\omega}^* < \tilde{m}^*$ for all $\tilde{\omega}^* < 1$, we must have $V''(z) \geq 0$ for all $z < 1$. □

Lemma 6. *Let $\omega^* \in [0,1]$ be a maximum of W. Define*

$$\bar{V}(m) = \begin{cases} V(m), & m < \omega^* \\ V\big(m^*\big) + V'\big(m^*\big)\big(m-m^*\big), & m \geq \omega^*. \end{cases} \tag{17}$$

If V is S-shaped on $[0, 1]$*, then* $\bar{V}$ *is convex on* $[0, 1]$ *and satisfies* $\bar{V}(m) \geq V(m)$ *for all* $m \in [0, 1]$*. If, in addition,* V *is strictly concave on* $[m^* - \varepsilon, 1]$ *for some* $\varepsilon > 0$*, then* $\bar{V}(m) > V(m)$ *for all* $m \in (\omega^*, m^*) \cup (m^*, 1]$*.*

Proof. Suppose that V is S-shaped on $[0, 1]$. As in Lemma 4, we can write

$$D(m) = V(m^*) + V'(m^*)(m - m^*) - V(m) = \int_m^{m^*} \left(-V''(z)\right)(z - m)\, \mathrm{d}z.$$

By Lemma 2, we have $\omega^* \leq x \leq m^*$. Moreover, by Lemma 4, we have $D(\omega^*) \geq 0$. Then, by (10) in Lemma 5 with $\omega_1 = \omega^*$, $\omega_2 = m$, and $m_1 = m_2 = m^*$, we have $D(m) \geq 0$ for $m \geq \omega^*$. Hence, since $V''(m) \geq 0$ for $m < \omega^*$, it follows that $\bar{V}$ is convex on $[0, 1]$ and satisfies $\bar{V}(m) \geq V(m)$ for all $m \in [0, 1]$. If, in addition, V is strictly concave on $[m^* - \varepsilon, 1]$, then $\bar{V}(m) > V(m)$ for all $m \in (\omega^*, m^*) \cup (m^*, 1]$, by (11) in Lemma 5. □

Lemma 7. *Let* $\bar{V}$ *be given by (17). If* $\bar{V}$ *is convex on* $[0, 1]$ *and satisfies* $\bar{V}(m) \geq V(m)$ *for all* $m \in [0, 1]$*, then upper censorship with cutoff* ω^* *is optimal.*

If, in addition, $\bar{V}$ *is strictly convex on* $[0, \omega^*]$ *and satisfies* $\bar{V}(m) > V(m)$ *for all* $m \in (\omega^*, m^*) \cup (m^*, 1]$*, then upper censorship with cutoff* ω^* *is uniquely optimal.*

Proof. The optimality of upper censorship with cutoff ω^* follows from Proposition 3 in Kolotilin (2018). For completeness, we include a simple self-contained proof, inspired by the proof of Theorem 1 in Dworczak and Martini (2019). Consider an arbitrary signal s. Let H be the distribution of $m = \mathbb{E}[\omega|s]$ induced by signal s. The sender's expected utility is smaller under signal s than under upper censorship with cutoff ω^* because

$$\int_0^1 V(m)\, \mathrm{d}H(m) \leq \int_0^1 \bar{V}(m)\, \mathrm{d}H(m) \tag{18}$$

$$\leq \int_0^1 \bar{V}(m)\, \mathrm{d}F(m) \tag{19}$$

$$= \int_0^{\omega^*} V(m)\, \mathrm{d}F(m) + \int_{\omega^*}^1 \bar{V}(m)\, \mathrm{d}F(m) \tag{20}$$

$$= \int_0^{\omega^*} V(m)\, \mathrm{d}F(m) + \int_{\omega^*}^1 V(m^*)\, \mathrm{d}F(m), \tag{21}$$

where (18) holds because $\bar{V}(m) \geq V(m)$ for all $m \in [0, 1]$, (19) holds because $\bar{V}$ is convex on $[0, 1]$ and F is a mean-preserving spread of H, (20) holds because $\bar{V}$ is given by (17), and (21) holds because $\bar{V}$ is linear in m on $[\omega^*, 1]$.

The unique optimality of upper censorship with cutoff ω^* can be shown using Proposition 2 in Dworczak and Martini (2019). To prove this directly, we first notice that the inequalities (18) and (19) must hold with equality for each optimal distribution H. Since V is convex on $[0, 1]$ and strictly convex on $[0, \omega^*]$, the inequality (19) can hold with equality only if the states below ω^* are revealed. Since $\bar{V}(m) > V(m)$ for all $m \in (\omega^*, m^*) \cup (m^*, 1]$, the inequality (18) can hold with equality only if the remaining states $[\omega^*, 1]$ are pooled, thus inducing the expected state m^* of the pool such that $\bar{V}(m^*) = V(m^*)$. □

The next lemma, whose proof is due to Ju Hu, shows that if V is not S-shaped, then, for some density f, the sender's expected utility is strictly greater under some lower-censorship signal than under any upper-censorship signal. Let $W_U(\omega)$ and $W_L(\omega)$ be the sender's expected utility under upper and lower censorship with a cutoff ω, respectively:

$$W_U(\omega) = \int_0^{\omega} V(m)\,\mathrm{d}F(m) + (1 - F(\omega))V(\mathbb{E}[m|m \geq \omega]).$$

$$W_L(\omega) = F(\omega)V(\mathbb{E}[m|m \leq \omega]) + \int_{\omega}^{1} V(m)\,\mathrm{d}F(m).$$

Lemma 8. *Let V be not S-shaped on* $[0, 1]$. *Then there exists a strictly positive density f on* $[0, 1]$ *such that $W_L(\omega^\star) > W_U(\omega)$ for some $\omega^\star \in (0, 1)$ and all $\omega \in [0, 1]$.*

Proof. First, note that it suffices to establish the existence of a non-negative density f on $[0, 1]$ such that $W_L(\omega^\star) > W_U(\omega)$ for some $\omega^\star \in (0, 1)$ and all $\omega \in [0, 1]$, because, by the continuity of V on $[0, 1]$, the inequality $W_L(\omega^\star) > W_U(\omega)$ continues to hold under a strictly positive density $(1 - \varepsilon)f + \varepsilon$, defined on $[0, 1]$, for small enough $\varepsilon > 0$.

For each $m, \omega \in [0, 1]$, define $D(m, \omega) = V(m) + V'(m)(\omega - m) - V(\omega)$. Since V is not S-shaped, there exist $m_1 < m_2 \leq m_3 < m_4$ such that V is strictly concave on $[m_1, m_2]$ and strictly convex on $[m_3, m_4]$. We claim that it suffices to consider the two cases:

Case 1. $D(m, \omega) \geq 0$ for all $\omega \in [m_1, m_2]$ and all $m \in [m_3, m_4]$;

Case 2. $D(m, \omega) < 0$ for all $\omega \in [m_1, m_2]$ and all $m \in [m_3, m_4]$.

To see this, consider first the case where $D(m_4, \omega) \geq 0$ for all $\omega \in [m_1, m_2]$. Since V is convex on $[m_3, m_4]$, we clearly have $D(m, \omega) \geq D(m_4, \omega) \geq 0$ for all $\omega \in [m_1, m_2]$ and all $m \in [m_3, m_4]$, so Case 1 holds. Consider now the remaining case where $D(m_4, \omega) < 0$ for some $\omega \in [m_1, m_2]$. Since V is continuously differentiable on $[0, 1]$, there exist $m_1', m_2' \in [m_1, m_2]$ with $m_1' < m_2'$ and $m_3' \in (m_3, m_4)$ such that $D(m, \omega) < 0$ for all $\omega \in [m_1', m_2']$ and all $m \in [m_3', m_4]$, so Case 2 holds after redefining $m_1 = m_1'$, $m_2 = m_2'$, and $m_3 = m_3'$.

To prove the lemma, it remains to show that there exists small enough $\varepsilon > 0$ such that $W_L(m_2) > W_U(\omega)$ for all $\omega \in [0, 1]$ when the density f is given by

$$f(\omega) = \begin{cases} \dfrac{\varepsilon}{m_2 - m_1}, & m \in [m_1, m_2], \\ \dfrac{1 - \varepsilon}{m_4 - m_3}, & m \in [m_3, m_4], \\ 0, & m \notin [m_1, m_2] \cup [m_3, m_4]. \end{cases}$$

Since W_U is constant over the intervals $[0, m_1]$, $[m_2, m_3]$, $[m_4, 1]$, we only need to consider $\omega \in [m_1, m_2] \cup [m_3, m_4]$. For $\omega \in [m_3, m_4]$, we have

$$W_L(m_2) > W_L(m_1) = W_L(0) = W_U(1) = W_U(m_4) \geq W_U(\omega),$$

where the first inequality is by strict concavity of V on $[m_1, m_2]$, the last inequality is by convexity of V on $[m_3, m_4]$, and the equalities are by the definition of W_L, W_U, and f.

Finally, consider the interval $[m_1, m_2]$. There exists small enough $\varepsilon > 0$ such that $\mathbb{E}[m] > m_3$ and $\mathbb{E}[V(m)] > V(\mathbb{E}[m])$, where the latter inequality holds because V is strictly convex on $[m_3, m_4]$ and continuously differentiable on $[0, 1]$.

In Case 1, for $\omega \in [m_1, m_2]$, we have

$$\begin{aligned}
W_L(m_2) &\geq W_L(m_1) \\
&= W_U(\omega) + \int_\omega^1 V(m)\,\mathrm{d}F(m) - (1 - F(\omega))V(\mathbb{E}[m|m \geq \omega]) \\
&\geq W_U(\omega) + \int_{m_1}^1 V(m)\,\mathrm{d}F(m) - (1 - F(m_1))V(\mathbb{E}[m|m \geq m_1]) \\
&> W_U(\omega),
\end{aligned}$$

where the first inequality is by concavity of V on $[m_1, m_2]$, the equality is by the definition of W_L and W_U, the second inequality is by

$$\frac{\mathrm{d}}{\mathrm{d}\omega}\left(\int_\omega^1 V(m)\,\mathrm{d}F(m) - (1 - F(\omega))V(\mathbb{E}[m|m \geq \omega])\right) = f(\omega)D(\mathbb{E}[m|m \geq \omega], \omega) > 0,$$

which follows from $\mathbb{E}[m|m \geq \omega] \geq \mathbb{E}[m] > m_3$ and the condition of Case 1, and the last inequality is by $\mathbb{E}[V(m)] > V(\mathbb{E}[m])$.

In Case 2, for $\omega \in [m_1, m_2]$, we have

$$\begin{aligned}
W_L(m_2) &\geq W_L(m_1) \\
&= W_U(\omega) + \int_\omega^1 V(m)\,\mathrm{d}F(m) - (1 - F(\omega))V(\mathbb{E}[m|m \geq \omega]) \\
&\geq W_U(\omega) + \int_{m_2}^1 V(m)\,\mathrm{d}F(m) - (1 - F(m_2))V(\mathbb{E}[m|m \geq m_2]) \\
&= W_U(\omega) + \int_{m_3}^1 V(m)\,\mathrm{d}F(m) - (1 - F(m_3))V(\mathbb{E}[m|m \geq m_3]) \\
&> W_U(\omega),
\end{aligned}$$

where the first inequality is by concavity of V on $[m_1, m_2]$, the first equality is by the definition of W_L and W_U, the second inequality follows from $\mathbb{E}[m|m \geq \omega] \geq \mathbb{E}[m] > m_3$ and the condition of Case 2 by an argument analogous to that in Case 1, the second equality is by the definition of f, and the last inequality is by strict convexity of V on $[m_3, m_4]$. $\square$

Proof of Theorem 1. By Lemma 2, there exists $x \in [\omega^*, m^*]$ such that V is convex on $[0, x]$ and concave on $[x, 1]$. So Lemmas 6 and 7 imply that upper censorship with cutoff ω^* is (uniquely) optimal if V is (strictly) S-shaped.

Suppose that V is not S-shaped. Then, by Lemma 8, there exists a strictly positive density f and a cutoff $\omega^\star$ such that the sender's expected utility is strictly greater under the lower-censorship signal with cutoff $\omega^\star$ than under any upper-censorship signal.

Finally, suppose that V is S-shaped but not strictly S-shaped. Then there exists $m_1 < m_2$ such that $V''(z) = 0$ for $m \in [m_1, m_2]$. Moreover, there exist a strictly positive density f and cutoffs $\omega_1^* < \omega_2^*$ in the interval (m_1, m_2) such that $m_1^* = \mathbb{E}[\omega|\omega \geq \omega_1^*]$ and $m_2^* = \mathbb{E}[\omega|\omega \geq \omega_2^*]$ are also in (m_1, m_2). Then, by Lemma 7, the upper-censorship signals with the cutoffs ω_1^* and ω_2^* are both optimal. □

Proof of Corollary 1. If V is (strictly) convex/concave on $[0, 1]$, then W is (strictly) increasing/decreasing, by Lemma 4; so ω^* equal to $1/0$ is a (unique) maximum of W. Thus, full/no disclosure is (uniquely) optimal by Theorem 1. Conversely, if V is not convex/concave on $[0, 1]$, then there exists $m_1 < m_2$ such that V is strictly concave/convex on $[m_1, m_2]$. Then for any density f that is strictly positive only on $[m_1, m_2]$, no/full disclosure is uniquely optimal. Thus, for any strictly positive density sufficiently close to f, full/no disclosure is suboptimal. Finally, if V is convex/concave on $[0, 1]$ but not strictly so, then V is not strictly S-shaped, and thus upper censorship is not uniquely optimal by Theorem 1. □

Proof of Proposition 1. Let V_1 and V_2 be S-shaped on $[0, 1]$, and let V_2 be more risk averse than V_1. Define W_1 and W_2 by (1) for V_1 and V_2. It suffices to show that, for each $\omega \in [0, 1]$, $W_1'(\omega) \leq (<)0 \implies W_2'(\omega) \leq (<)0$.

Consider $\omega \in [0, 1]$ and $m = \mathbb{E}[\omega'|\omega' \geq \omega]$. We have

$$\begin{aligned} W_i'(\omega) &= f(\omega)\big(V_i(\omega) - V_i(m) + V_i'(m)(m-\omega)\big) \\ &= f(\omega)V_i'(m)\left(m - \omega - \frac{V_i(m) - V_i(\omega)}{V_i'(m)}\right) \\ &= f(\omega)V_i'(m)\left(m - \omega - \int_\omega^m \frac{V_i'(z)}{V_i'(m)}\,\mathrm{d}z\right). \end{aligned}$$

Since $f(\omega) > 0$ and $V_i'(m) > 0$, we obtain

$$W_i'(\omega) \leq (<)0 \quad \iff \quad m - \omega \leq (<) \int_\omega^m \frac{V_i'(z)}{V_i'(m)}\,\mathrm{d}z. \tag{22}$$

Next,

$$\int_\omega^m \frac{V_2'(z)}{V_2'(m)}\,\mathrm{d}z - \int_\omega^m \frac{V_1'(z)}{V_1'(m)}\,\mathrm{d}z = \int_\omega^m \frac{V_1'(z)}{V_2'(m)}\left(\frac{V_2'(z)}{V_1'(z)} - \frac{V_2'(m)}{V_1'(m)}\right)\mathrm{d}z \geq 0, \tag{23}$$

where the inequality holds because $V_2'(z)/V_1'(z)$ is decreasing in z, as follows from (4).

Finally, if $W_1'(\omega) \leq (<)0$, then $W_2'(\omega) \leq (<)0$, by (22) and (23). □

Proof of Proposition 2. Let V_1 be S-shaped on $\mathbb{R}$, and let V_2 be a (location–scale) shift to the left of V_1. Define W_1 and W_2 by (1) for V_1 and V_2. We first show that, for each $\omega \in [0, 1]$, $W_1'(\omega) \leq 0 \implies W_2'(\omega) \leq 0$.

Consider $\omega \in [0, 1]$ and $m = \mathbb{E}[\omega' | \omega' \geq \omega]$. We have

$$
\begin{aligned}
W_2'(\omega) &= f(\omega) \int_{\omega}^{m} V_2''(z)(z-\omega)\,\mathrm{d}z \\
&= \frac{f(\omega)}{\beta^2} \int_{\omega}^{m} V_1''\left(\frac{z-\alpha}{\beta}\right)(z-\omega)\,\mathrm{d}z \\
&= f(\omega) \int_{\frac{\omega-\alpha}{\beta}}^{\frac{m-\alpha}{\beta}} V_1''(y)\left(y - \frac{\omega-\alpha}{\beta}\right)\mathrm{d}y, \qquad (24)
\end{aligned}
$$

where the first line is by Lemma 4, the second line is by (5), and the third line is by the change of variable $y = (z-\alpha)/\beta$.

For $\omega = 1$, we have $m = 1$, so $W_2'(\omega) = W_1'(\omega) = 0$. For $\omega < 1$, we have $\omega < m < 1$. So if $W_1'(\omega) \leq 0$, then there exists $x_1 < m$ such that $V_1''(z) \geq (\leq) 0$ for $z < (>) x_1$, by Lemma 4 and S-shapedness of V_1. Moreover, $(\omega-\alpha)/\beta \geq \omega$ and $(m-\alpha)/\beta \geq m$, because $\alpha + \beta z \leq z$ for all $z \in [0, 1]$, by (5). Hence, if $W_1'(\omega) \leq 0$, then $W_2'(\omega) \leq 0$ by (10) in Lemma 5 with $\omega_1 = \omega$, $m_1 = m$, $\omega_2 = (\omega-\alpha)/\beta$, and $m_2 = (m-\alpha)/\beta$.

Let, for $z \in \mathbb{R}$ and $\delta \in (0, 1)$,

$$
V_\delta(z) = V_1\left(\delta \frac{z-\alpha}{\beta} + (1-\delta)z\right),
$$

and let W_δ be given by (1) for V_δ. Clearly, V_2 is a shift to the left of V_δ, and V_δ is a shift to the left of V_1. Thus, to complete the proof, it suffices to show that, for each $\omega \in [0, 1]$ such that $W_1'(\omega) < 0$, there exists $\delta \in (0, 1)$ such that $W_\delta'(\omega) < 0$. Since V_1 is S-shaped, $W_1'(\omega) < 0$, together with Lemma 4, implies that $\omega < x_1$, where x_1 is such that V_1 is convex on $(-\infty, x_1]$ and concave on $[x_1, \infty)$. Thus, there exists a sufficiently small $\delta > 0$ such that

$$
\delta \frac{\omega-\alpha}{\beta} + (1-\delta)\omega < x_1.
$$

Next, since, by Lemma 4, the inequality in (15) is strict when $W_1'(\omega) < 0$, it follows from (12)–(15) that

$$
\int_{\delta\frac{\omega-\alpha}{\beta}+(1-\delta)\omega}^{\delta\frac{m-\alpha}{\beta}+\delta m} V_1''(z)\left(z - \delta\frac{\omega-\alpha}{\beta} - (1-\delta)\omega\right)\mathrm{d}z < 0.
$$

Hence, applying (24) for V_δ, we obtain that $W_\delta'(\omega) < 0$. □

Proof of Corollary 2. Suppose that $g'(r)/g(r) = -\lambda$. By (8), V is strictly convex (concave) on $[0, 1]$ if $\rho > \lambda$ ($\rho < \lambda$), so full disclosure (no disclosure) is uniquely optimal by Corollary 1. If $\rho = \lambda$, then V is linear in m on $[0, 1]$, so all signals are optimal, because $\int_0^1 V(m)\,\mathrm{d}H(m) = \int_0^1 V(m)\,\mathrm{d}F(m)$ for all H such that F is a mean-preserving spread of H.

Conversely, if g is not exponential, then there exists $m_1 < m_2$ such that g'/g is strictly monotone on $[m_1, m_2]$. Suppose, for concreteness, that g'/g is strictly decreasing on $[m_1, m_2]$. Choosing $\rho = -g'(x)/g(x)$ for some $x \in (m_1, m_2)$, we obtain that V is strictly convex on $[m_1, x]$ and strictly concave on $[x, m_2]$. Then there exist $\omega^* \in (m_1, x)$, $m^* \in (x, m_2)$, and a density f that is strictly positive only on $[m_1, m_2]$ such that $m^* = \mathbb{E}[\omega | \omega \geq$

$\omega^*]$ and $\int_{\omega^*}^{m^*} V''(z)(z-\omega^*)\,\mathrm{d}z = 0$. Since V is strictly S-shaped, Lemma 1 implies that W is strictly quasi-concave; so ω^* is a maximum of W by Lemma 4. Consequently, the upper censorship signal with the interior cutoff ω^* is uniquely optimal by Theorem 1, and, thus, full and no disclosure are both suboptimal. Moreover, for any strictly positive density sufficiently close to f, full and no disclosure remain suboptimal. □

Proof of Proposition 3. Part (i). Let V_1 and V_2 be given by (6) with $\rho_1 > \rho_2$. We have

$$\begin{aligned} &V_1''(z)V_2'(z) - V_2''(z)V_1'(z) \\ &\quad = \big(g'(z) + \rho_1 g(z)\big)\big(g(z) + \rho_2 G(z)\big) - \big(g'(z) + \rho_2 g(z)\big)\big(g(z) + \rho_1 G(z)\big) \\ &\quad = (\rho_1 - \rho_2)\big(\big(g(z)\big)^2 - g'(z)G(z)\big) \ge 0, \end{aligned}$$

where the inequality holds because $\rho_1 > \rho_2$ and g is log-concave, implying that G is also log-concave (Theorem 1 in Bagnoli and Bergstrom (2005)), so $(g(z))^2 \ge g'(z)G(z)$. We thus obtain that V_2 is more risk averse than V_1. Consequently, by Proposition 1, the sender reveals less information under V_2 than under V_1.

Part (ii). Let G_2 be a (location–scale) shift to the left of G_1. Let, for $m \in \mathbb{R}$ and $i = 1, 2$,

$$V_i(m) = G_i(m) + \rho \int_{\infty}^{m} G_i(r)\,\mathrm{d}r,$$

and let, for $m \in \mathbb{R}$,

$$\tilde{V}_1(m) = G_1(m) + \rho\beta \int_{-\infty}^{m} G_1(r)\,\mathrm{d}r.$$

So $\tilde{V}_1$ is the same as V_1 except it has a smaller alignment parameter, $\rho\beta$. We have

$$\begin{aligned} V_2(m) &= G_1\left(\frac{m-\alpha}{\beta}\right) + \rho \int_{-\infty}^{m} G_1\left(\frac{r-\alpha}{\beta}\right)\mathrm{d}r \\ &= G_1\left(\frac{m-\alpha}{\beta}\right) + \rho\beta \int_{-\infty}^{\frac{m-\alpha}{\beta}} G_1(z)\,\mathrm{d}z = \tilde{V}_1\left(\frac{m-\alpha}{\beta}\right), \end{aligned}$$

where the second line is by the change of variable $z = (r-\alpha)/\beta$. That is, V_2 is a shift to the left of $\tilde{V}_1$, and $\tilde{V}_1$ is more risk averse than V_1, by part (i). Consequently, by Propositions 1 and 2, the sender reveals less information under V_2 than under V_1. □

Proof of Theorem 3. Let b and g be (strictly) log-concave. By Theorem 1, it suffices to show that V is (strictly) S-shaped on $[0, 1]$. By (9) with $\rho = 0$, we have

$$\begin{aligned} V''(m) &= \frac{\mathrm{d}^2}{\mathrm{d}m^2} B\big(G(m)\big) = b'\big(G(m)\big)\big(g(m)\big)^2 + b\big(G(m)\big)g'(m) \\ &= b\big(G(m)\big)\big(g(m)\big)^2\left(\frac{b'\big(G(m)\big)}{b\big(G(m)\big)} + \frac{g'(m)}{\big(g(m)\big)^2}\right). \end{aligned}$$

By assumption, b and g are strictly positive, so $b(G(m))(g(m))^2 > 0$. Because b is (strictly) log-concave and G is strictly increasing, $b'(G(m))/b(G(m))$ is (strictly) decreasing. Because g is log-concave, we have $g''(m)g(m) \leq (g'(m))^2$. Therefore,

$$\frac{\mathrm{d}}{\mathrm{d}m}\left(\frac{g'(m)}{\big(g(m)\big)^2}\right) = \frac{g''(m)\big(g(m)\big)^2 - 2g(m)\big(g'(m)\big)^2}{\big(g(m)\big)^4} \leq \frac{\big(g'(m)\big)^2 g(m) - 2g(m)\big(g'(m)\big)^2}{\big(g(m)\big)^4}$$

$$= -\frac{\big(g'(m)\big)^2}{\big(g(m)\big)^3} \leq 0.$$

Thus, g'/g^2 is decreasing. We have proved that V is (strictly) S-shaped on $[0, 1]$. □

References

Alonso, Ricardo and Odilon Câmara (2016a), "Persuading voters." *American Economic Review*, 3590–3605. [563]

Alonso, Ricardo and Odilon Câmara (2016b), "Political disagreement and information in elections." *Games and Economic Behavior*, 100, 390–412. [563]

Arieli, Itai, Yakov Babichenko, Rann Smorodinsky, and Takuro Yamashita (2021), "Optimal persuasion via bi-pooling." Report. [563]

Bagnoli, Mark and Ted Bergstrom (2005), "Log-concave probability and its applications." *Economic Theory*, 08, 445–469. [570, 582]

Bergemann, Dirk and Stephen Morris (2019), "Information design: A unified perspective." *Journal of Economic Literature*, 57, 44–95. [561]

Besley, Timothy and Andrea Prat (2006), "Handcuffs for the grabbing hand? Media capture and government accountability." *American Economic Review*, 96, 720–736. [574]

Bloedel, Alexander W. and Ilya Segal (2021), "Persuading a rationally inattentive agent." Report. [563]

Chan, Jimmy and Wing Suen (2008), "A spatial theory of news consumption and electoral competition." *Review of Economic Studies*, 75, 699–728. [574]

Chiang, Chun-Fang and Brian Knight (2011), "Media bias and influence: Evidence from newspaper endorsements." *Review of Economic Studies*, 795–820. [574]

Dizdar, Deniz and Eugen Kováč (2020), "A simple proof of strong duality in the linear persuasion problem." *Games and Economic Behavior*, 122, 407–412. [563]

Duffie, Darrell, Piotr Dworczak, and Haoxiang Zhu (2017), "Benchmarks in search markets." *Journal of Finance*, 72, 1983–2044. [563]

Dworczak, Piotr and Giorgio Martini (2019), "The simple economics of optimal persuasion." *Journal of Political Economy*, 993–2048. [563, 577]

Edmond, Chris (2013), "Information manipulation, coordination, and regime change." *Review of Economic Studies*, 1422–1458. [574]

Egorov, Georgy, Sergei Guriev, and Konstantin Sonin (2009), "Why resource-poor dictators allow freer media: A theory and evidence from panel data." *American Political Science Review*, 103, 645–668. [574]

Gehlbach, Scott and Konstantin Sonin (2014), "Government control of the media." *Journal of Public Economics*, 118, 163–171. [563, 572, 574]

Gentzkow, Matthew and Emir Kamenica (2016), "A Rothschild-Stiglitz approach to Bayesian persuasion." *American Economic Review: Papers & Proceedings*, 597–601. [563]

Ginzburg, Boris (2019), "Optimal information censorship." *Journal of Economic Behavior and Organization*, 163, 377–385. [563]

Gitmez, Arda and Pooya Molavi (2020), "Media capture: A Bayesian persuasion approach." Report. [563, 572]

Goldstein, Itay and Yaron Leitner (2018), "Stress tests and information disclosure." *Journal of Economic Theory*, 34–69. [563]

Guo, Yingni and Eran Shmaya (2019), "The interval structure of optimal disclosure." *Econometrica*, 87, 653–675. [563]

Johnson, Justin P. and David P. Myatt (2006), "On the simple economics of advertising, marketing, and product design." *American Economic Review*, 96, 756–784. [571]

Kamenica, Emir (2019), "Bayesian persuasion and information design." *Annual Review of Economics*, 11, 249–272. [561]

Kamenica, Emir and Matthew Gentzkow (2011), "Bayesian persuasion." *American Economic Review*, 101, 2590–2615. [561, 563]

Kleiner, Andreas, Benny Moldovanu, and Philipp Strack (2021), "Extreme points and majorization: Economic applications." *Econometrica*, 89, 1557–1593. [563]

Kolotilin, Anton (2015), "Experimental design to persuade." *Games and Economic Behavior*, 90, 215–226. [563]

Kolotilin, Anton (2018), "Optimal information disclosure: A linear programming approach." *Theoretical Economics*, 13, 607–636. [563, 564, 577]

Kolotilin, Anton and Hongyi Li (2021), "Relational communication." *Theoretical Economics*, 16, 1391–1430. [563]

Kolotilin, Anton, Ming Li, Tymofiy Mylovanov, and Andriy Zapechelnyuk (2015), "Persuasion of a privately informed receiver." Report. [561]

Kolotilin, Anton, Tymofiy Mylovanov, Andriy Zapechelnyuk, and Ming Li (2017), "Persuasion of a privately informed receiver." *Econometrica*, 85, 1949–1964. [561, 563, 574]

Kolotilin, Anton and Alexander Wolitzky (2020), "Assortative information disclosure." Report. [563]

Lewis, Tracy R. and David Sappington (1994), "Supplying information to facilitate price discrimination." *International Economic Review*, 35, 309–327. [571]

Lipnowski, Elliot, Laurent Mathevet, and Dong Wei (2020), "Attention management." *American Economic Review: Insights*, 17–32. [563]

Lorentzen, Peter (2014), "China's strategic censorship." *American Journal of Political Science*, 402–414. [574]

Milgrom, Paul R. (1981), "Good news and bad news: Representation theorems and applications." *Bell Journal of Economics*, 12, 380–391. [572]

Orlov, Dmitry, Pavel Zryumov, and Andrzej Skrzypach (2021), "Design of macro-prudential stress tests." Report. [563]

Ostrovsky, Michael and Michael Schwarz (2010), "Information disclosure and unraveling in matching markets." *American Economic Journal: Microeconomics*, 34–63. [563]

Prat, Andrea and David Strömberg (2013), "The political economy of mass media." *Advances in Economics and Econometrics*, 135–187. [574]

Rayo, Luis and Ilya Segal (2010), "Optimal information disclosure." *Journal of Political Economy*, 949–987. [561, 563]

Romanyuk, Gleb and Alex Smolin (2019), "Cream skimming and information design in matching markets." *American Economic Journal: Microeconomics*, 250–276. [563]

Suen, Wing (2004), "The self-perpetuation of biased beliefs." *Economic Journal*, 114, 377–396. [574]

Zapechelnyuk, Andriy (2020), "Optimal quality certification." *American Economic Review: Insights*, 161–176. [563]

Co-editor Simon Board handled this manuscript.

Manuscript received 10 December, 2019; final version accepted 29 May, 2021; available online 15 June, 2021.