ON MONOTONE PERSUASION

ANTON KOLOTILIN
School of Economics, UNSW Business School

HONGYI LI
School of Economics, UNSW Business School

ANDRIY ZAPECHELNYUK
School of Economics, University of Edinburgh

We study monotone persuasion in the linear case, where a posterior distribution over states is summarized by its mean. We identify two settings where the optimal unrestricted signal can be nonmonotone. In the first setting, the optimal unrestricted signal requires randomization. In the second setting, the optimal unrestricted signal entails nonmonotone pooling of states. We solve for the optimal monotone signal in each setting, and illustrate our results with an application to media censorship.

KEYWORDS: Bayesian persuasion, monotone persuasion.
JEL CLASSIFICATION. D82, D83.

1. INTRODUCTION

The literature on Bayesian persuasion has largely focused on the *linear* case, where the state space is one-dimensional and a posterior distribution over states is summarized by its mean (e.g., Gentzkow and Kamenica 2016, Kolotilin et al. 2017, Kolotilin 2018, Dworczak and Martini 2019). The standard approach has been to analyze *unrestricted* persuasion, where the set of feasible signals is unrestricted. In reality, however, various feasibility constraints bind due to incentive, legal, or other practical reasons. One such constraint is that signals should be *deterministic*. For instance, the bank regulator may be precluded from implementing a stochastic stress test if they cannot credibly and verifiably randomize scores. A second constraint is that signals should be *monotone*. For instance, a bank regulator may be precluded from using a stress test that gives a higher score to a weaker bank.

These concerns have led to a literature on *monotone persuasion* (Ivanov 2021, Mensch 2021, and Onuchic and Ray 2023) where all feasible signals are deterministic and monotone in that they partition the state space into convex sets (i.e., intervals and singletons).[1] However, these papers do not address the linear case. In the literature that deals with the linear case, Dworczak and Martini (2019) delineate conditions under which monotone persuasion is optimal, so that standard results from unrestricted persuasion apply.[2] It remains an open question what

Anton Kolotilin: akolotilin@gmail.com
Hongyi Li: hongyi@gmail.com
Andriy Zapechelnyuk: azapech@gmail.com

Early versions of the results in this paper were presented in working paper versions of Kolotilin and Li (2021) and Kolotilin et al. (2022). We thank Simon Board and Tymofiy Mylovanov for helpful comments. Kolotilin gratefully acknowledges financial support from the Australian Research Council Discovery Early Career Research Award DE160100964. Zapechelnyuk gratefully acknowledges financial support from the Economic and Social Research Council Grant ES/N01829X/1.

[1] Relatedly, Aybas and Turkel (2024), Hopenhayn and Saeedi (2024), and Lyu et al. (2024) study Bayesian persuasion problems under a different feasibility constraint: the set of signal realizations contains only N elements.

[2] Some other restrictions are known to be innocuous in the linear case. In particular, stochastic monotone signals – where a higher state induces a higher lottery over signal realizations in the likelihood ratio order – are without loss of

the solution to the monotone persuasion problem is when optimal unrestricted signals are (i) non-deterministic or (ii) nonmonotone. We take first steps towards answering this question by deriving the optimal monotone signal in two distinct settings which separately address (i) and (ii).

First, we consider the simplest setting where randomization is valuable: the state is discrete, and the objective function is s-shaped (convex, then concave). Here, it is known that the optimal unrestricted signal involves "stochastic upper censorship" where low states are separated, high states are pooled, and the cutoff state is randomly either separated or pooled with high states. We show that the optimal monotone signal has the same "upper censorship" form, but does not randomize at the cutoff state.

Second, we consider the simplest setting where nonmonotone pooling of states is valuable: the state is continuous, and the objective function is m-shaped (concave-convex-concave). If an optimal unrestricted signal is nonmonotone, then it induces two signal realizations that concavify the objective function. In this case, we show that the optimal monotone signal partitions the state space into one or two intervals.

Our two settings are chosen for simplicity, but the methodology we develop can be adapted to analyse more general cases. In particular, it is straightforward to allow for (i) a convex-concave-convex objective if the state is discrete and (ii) a convex-concave-convex-concave-convex objective if the state is continuous. In other words, adding a convex region on any side of the objective does not present any further conceptual challenges. More general cases where the objective function has even more inflection points are less likely to be relevant in applications.

We illustrate our results using Kolotilin et al.'s (2022) media censorship model, which features a government, heterogeneous citizens, and media outlets. We show that the government's problem reduces to a monotone persuasion problem. Our first setting corresponds to the case where there is initially a finite number of media outlets and the distribution of citizens' types is unimodal. In this case, the government permits all sufficiently supportive media outlets and censors all other media outlets, which extends Kolotilin et al.'s (2022) result on the optimality of upper censorship from the continuous case to the discrete one. Our second setting corresponds to the case where there is initially a continuum of media outlets and the distribution of citizens' types is bimodal (i.e., society is polarized). In this case, the government permits at most one media outlet and censors all other media outlets.

2. MODEL

A *state* $\omega \in [0,1]$ is a random variable with a prior probability distribution function F. A *signal* reveals information about the state. An *objective* $V : [0,1] \mapsto \mathbb{R}$ is a twice continuously differentiable function of the expected state m induced by a signal.

In many applications, the state is either continuous or discrete. The state is *continuous* if F has a strictly positive density f on $[0,1]$. The state is *discrete* if the support of F, denoted by $\text{supp}(F)$, is a finite subset of $[0,1]$. The discrete density is also denoted by f.

In an *unrestricted persuasion problem*, a signal can be arbitrarily correlated with the state. By Blackwell (1951), there exists a signal that induces a probability distribution G of the expected state m iff the prior distribution F is a mean-preserving spread of G (e.g., Kolotilin 2018).

generality (Kolotilin and Zapechelnyuk 2024), while deterministic (possibly nonmonotone) signals are without loss of generality if the state is continuous (Arieli et al. 2023).

Thus, the unrestricted persuasion problem is to maximize $\int_0^1 V(m)\mathrm{d}G(m)$ over distributions G such that F is a mean-preserving spread of G.

In a *monotone persuasion problem*, a signal is required to be *monotone*: it pools the states into convex sets (i.e., intervals and singletons) and reveals which set contains the realized state. Formally, a monotone signal is an increasing function $\mu : [0,1] \mapsto [0,1]$. W.l.o.g., we identify each realization m with the expected state induced by this realization, so $m = \mathbb{E}[\omega|\mu(\omega) = m]$. Let $\mathcal{M}$ be the set of monotone signals. Thus, the monotone persuasion problem is to maximize $\int_0^1 V(\mu(\omega))\mathrm{d}F(\omega)$ over monotone signals $\mu \in \mathcal{M}$.

Dworczak and Martini (2019) show that there exists a monotone signal that solves the unrestricted persuasion problem if the state is continuous and the objective function is *affine closed*. In particular, V is affine closed if it has no m-shaped region: there do not exist $0 \le \underline{\omega} < \omega_L < \omega_R < \overline{\omega} \le 1$ such that V is strictly concave on $[\underline{\omega}, \omega_L]$, strictly convex on $[\omega_L, \omega_R]$, and strictly concave on $[\omega_R, \overline{\omega}]$.

We study two simplest cases where no monotone signal solves the unrestricted persuasion problem. In Section 3, the objective is s-shaped but the state is discrete. In Section 4, the state is continuous but the objective is m-shaped. In both cases, existing approaches from the Bayesian persuasion literature, such as concavification and linear programming duality, no longer apply, because the monotone persuasion problem is not a linear program.

3. DISCRETE STATE AND S-SHAPED OBJECTIVE

In this section, the state is discrete and the objective is s-shaped. The objective function V is *s-shaped* if there exists $0 < \omega_M < 1$ such that V is strictly convex on $[0, \omega_M]$ and strictly concave on $[\omega_M, 1]$.

A signal is *stochastic upper censorship* if there exist $\omega^* \in \operatorname{supp}(F)$ and $q^* \in [0,1]$ such that states in $[0, \omega^*)$ are separated, states in $(\omega^*, 1]$ are pooled, and state ω^* is separated with probability q^* and pooled with probability $1 - q^*$. Let

$$m^* = \frac{\omega^*(1-q^*)f(\omega^*) + \sum_{\omega > \omega^*} \omega f(\omega)}{(1-q^*)f(\omega^*) + \sum_{\omega > \omega^*} f(\omega)} \tag{1}$$

be the expected state conditional on the pooling signal realization. A stochastic upper-censorship signal with (ω^*, q^*) is *deterministic upper censorship* if $q^* \in \{0,1\}$. This is the monotone signal μ given by

$$\mu(\omega) = \begin{cases} \omega, & \omega \in [0, \omega^*), \\ \omega^*, & \omega = \omega^* \text{ and } q^* = 1, \\ m^*, & \omega = \omega^* \text{ and } q^* = 0, \\ m^*, & \omega \in (\omega^*, 1], \end{cases}$$

By Alonso and Câmara (2016) and Kolotilin et al. (2022), there exist $\omega^* \in \operatorname{supp}(F)$ and $q^* \in [0,1]$ satisfying

$$V(m^*) + V'(m^*)(\omega^* - m^*) \ge V(\omega^*), \quad \text{with equality if } (\omega^*, q^*) \ne (0,0), \tag{2}$$

such that stochastic upper censorship with (ω^*, q^*) solves the unrestricted persuasion problem.

THEOREM 1: *If the state is discrete, V is s-shaped, and (ω^*, q^*) is given by (2), then there exists $q^{**} \in \{0,1\}$ such that deterministic upper censorship with (ω^*, q^{**}) solves the monotone persuasion problem.*

One may expect that if all deterministic (not necessarily monotone) signals were allowed, then deterministic upper censorship would still be optimal when V is s-shaped. This, however, is not true. For example, suppose that the state ω takes values 0, ε, and 1 with probabilities $1/4$, $1/2$, and $1/4$, where $\varepsilon > 0$ is sufficiently small. Suppose that V is such that stochastic upper censorship that separates state ε with probability $q^* = 1/2$ solves the unrestricted persuasion problem. In this case, the optimal deterministic signal would pool states 0 and 1, and separate state ε.

4. CONTINUOUS STATE AND M-SHAPED OBJECTIVE

In this section, the state is continuous and the objective is m-shaped. The objective function V is *m-shaped* if there exist $0 < \omega_L < \omega_R < 1$ such that V is strictly concave on $[0, \omega_L]$, strictly convex on $[\omega_L, \omega_R]$, and strictly concave on $[\omega_R, 1]$.

A monotone signal μ is *interval disclosure* with cutoffs $0 \le \omega_L^* \le \omega_R^* \le 1$ if it separates states in the middle interval $[\omega_L^*, \omega_R^*]$ and pools states in the left interval $[0, \omega_L^*)$ and in the right interval $(\omega_R^*, 1]$, so

$$\mu(\omega) = \begin{cases} m_L^*, & \omega \in [0, \omega_L^*), \\ \omega, & \omega \in [\omega_L^*, \omega_R^*], \\ m_R^*, & \omega \in (\omega_R^*, 1], \end{cases}$$

where

$$m_L^* = \mathbb{E}\big[\omega | \omega \in [0, \omega_L^*]\big] \quad \text{and} \quad m_R^* = \mathbb{E}\big[\omega | \omega \in [\omega_R^*, 1]\big]$$

are the expected states conditional on the pooling signal realizations. A monotone signal μ is a *cutoff rule* with cutoff ω^* if it pools states in the intervals $[0, \omega^*)$ and $(\omega^*, 1]$. Finally, a monotone signal μ is *no disclosure* if it pools all states in $[0, 1]$. Note that no disclosure is a special case of a cutoff rule, which is, in turn, a special case of interval disclosure.

THEOREM 2: *If the state is continuous and V is m-shaped, then interval disclosure solves the monotone persuasion problem, as follows:*

(i) *If there exist $\omega_L^*, \omega_R^* \in (\omega_L, \omega_R)$ with $\omega_L^* < \omega_R^*$ such that*

$$V(m_L^*) + V'(m_L^*)(\omega_L^* - m_L^*) = V(\omega_L^*), \tag{3}$$

$$V(m_R^*) + V'(m_R^*)(\omega_R^* - m_R^*) = V(\omega_R^*), \tag{4}$$

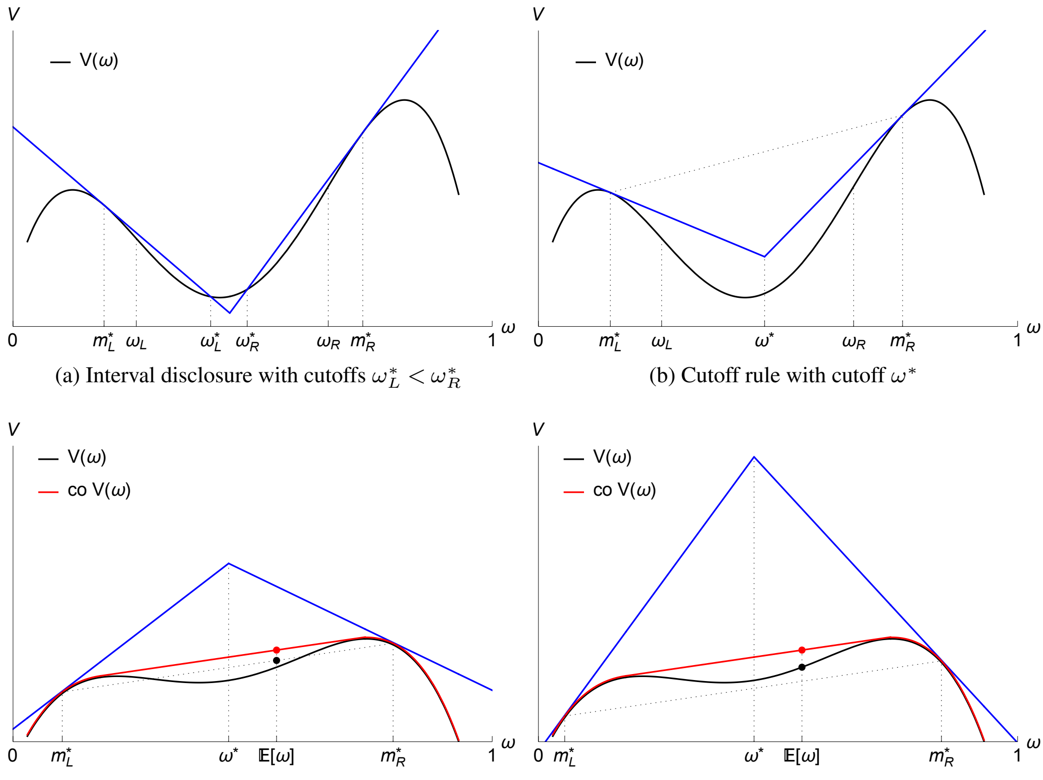
then interval disclosure with cutoffs ω_L^* and ω_R^* solves the monotone persuasion problem. Also, $m_L^* \in (0, \omega_L)$ and $m_R^* \in (\omega_R, 1)$.

(ii) *Else if there exists $\omega^* = \omega_L^* = \omega_R^* \in (0, 1)$ such that*

$$V(m_L^*) + V'(m_L^*)(\omega^* - m_L^*) = V(m_R^*) + V'(m_R^*)(\omega^* - m_R^*), \tag{5}$$

$$V(m_L^*)F(\omega^*) + V(m_R^*)(1 - F(\omega^*)) \ge V(\mathbb{E}[\omega]), \tag{6}$$

then the cutoff rule with cutoff ω^* solves the monotone persuasion problem. Also, $m_L^* \in (0, \omega_L)$ and $m_R^* \in (\omega_R, 1)$.



(a) Interval disclosure with cutoffs $\omega_L^* < \omega_R^*$

(b) Cutoff rule with cutoff ω^*

(c) Cutoff rule with cutoff ω^*

(d) No disclosure

FIGURE 1.—Interval disclosure when V is m-shaped

(iii) *Else, no disclosure is optimal.*

Theorem 2 shows that a signal that solves the monotone persuasion problem takes one of three forms: interval disclosure (Figure 1a), a cutoff rule (Figures 1b and 1c), and no disclosure (Figure 1d). Moving along Figures 1a → 1b → 1c → 1d, the prior distribution F puts increasingly more weight on left and right states (and less weight on middle states).

If (3) and (4) hold (Figure 1a), or if (5) and (6) hold and $V'(m_L^*) \leq V'(m_R^*)$ (Figure 1b), then optimal interval disclosure solves the unrestricted persuasion problem (see Kolotilin 2018, Proposition 3).

Otherwise (Figures 1c and 1d), there is a continuum of distinct signals that solve the unrestricted persuasion problem. All of them are nonmonotone and induce the same two expected states that yield the value $\operatorname{co} V(\mathbb{E}[\omega])$, where $\operatorname{co} V(\mathbb{E}[\omega])$ is the concavification of V at $\mathbb{E}[\omega]$, shown by the red dot in Figures 1c and 1d.[3] In contrast, the signal described in parts (ii) and (iii) of Theorem 2 uniquely solves the monotone persuasion problem and yields a strictly lower value, shown by the black dot in Figures 1c and 1d.

[3]The concavification of V is defined as $\operatorname{co} V(\omega) = \min_{v \in \mathcal{V}} v(\omega)$, for $\omega \in [0,1]$, where $\mathcal{V}$ is the set of all concave functions v on $[0,1]$ such that $v(\omega) \geq V(\omega)$ for all $\omega \in [0,1]$.

5. APPLICATION TO MEDIA CENSORSHIP

We illustrate our results using a simplified version of the media censorship model in Kolotilin et al. (2022). There is a government and a continuum of heterogeneous citizens. The government's quality $\theta \in [0,1]$ has a distribution T with a strictly positive density on $[0,1]$. Citizens are indexed by $r \in [0,1]$ that has a distribution V with a continuously differentiable density on $[0,1]$. The utility of a citizen of type r is

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

where $a_r \in \{0,1\}$ is the citizen's action. The government's utility is the aggregate action in the society $\int_0^1 a_r \mathrm{d}V(r)$.

Citizens obtain information about the government's quality 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 endorses action $a = 0$ if $\theta < c$. The set of media outlets C is a subset of $[0,1]$.

The government's censorship policy is a set of media outlets $X \subset C$ that are censored. The other media outlets in $C \backslash X$ are permitted to broadcast.

The timing is as follows. First, the government chooses a set $X \subset C$ of censored media outlets. Second, the government's quality θ is realized, and each permitted media outlet endorses action $a = 1$ or $a = 0$ according to its editorial policy. Finally, each citizen observes messages from all permitted media outlets, updates beliefs about θ, and chooses an action.

Consider a censorship policy $X \subset C$. Let y_X be a random variable equal to the conditional expectation of θ given messages from all media outlets in $C \backslash X$. Let G_X denote the distribution of y_X. Each citizen of type r chooses $a_r = 1$ iff $r \leq y_X$. Then, the aggregate action is $\int_0^1 a_r \mathrm{d}V(r) = V(y_X)$, and the government's expected utility is $\int_0^1 V(y)\mathrm{d}G_X(y)$. Let $\mathcal{G}_C$ denote the set of distributions G_X induced by all censorship policies $X \subset C$.

Define the state ω as the conditional expectation of θ given messages from all media outlets in C. That is, $\omega = y_\varnothing$ and its distribution is $F = G_\varnothing$. Consider a monotone signal μ, which is an increasing function satisfying $\mathbb{E}[\omega|\mu(\omega) = m] = m$ for all m. Let G_μ denote the distribution of $m = \mu(\omega)$. Then, the value of μ is $\int_0^1 V(\mu(\omega))\mathrm{d}F(\omega) = \int_0^1 V(m)\mathrm{d}G_\mu(m)$. Let $\mathcal{G}_M$ denote the set of distributions G_μ induced by all monotone signals $\mu \in \mathcal{M}$.

The next proposition shows that an outcome is implementable by a monotone signal iff it is implementable by a censorship policy. Thus, the media censorship problem reduces to a monotone persuasion problem.

PROPOSITION 1: $\mathcal{G}_C = \mathcal{G}_M$.

To see the intuition behind Proposition 1, suppose that there is only one media outlet with an editorial policy $c \in (0,1)$. There are two states, $\mathbb{E}[\theta|\theta \leq c]$ and $\mathbb{E}[\theta|\theta \geq c]$. Every monotone signal induces one of two possible outcomes: the states are separated or pooled. Each of the two outcomes is implementable by permitting or censoring the media outlet.

If there is a finite number of media outlets, then the state ω is discrete. By Theorem 1, if V is s-shaped (i.e., the distribution of citizens' types is unimodal), then the government optimally censors all media outlets whose editorial policies are below some cutoff. That is, all censored media outlets are less supportive than all permitted media outlets in that they endorse the government's preferred action $a = 1$ less frequently.

If there is a continuum of media outlets $C = [0,1]$, then the state ω is continuous and has distribution $F = T$. By Theorem 2, if V is m-shaped (i.e., the distribution of citizens' types is bimodal), then the government's optimal censorship policy takes one of three forms.[4] At one extreme, if the distribution of the government's quality T is sufficiently concentrated (part (i) of Theorem 2), then the government optimally permits a range of moderate media outlets. At the other extreme, if T is sufficiently spread out (part (iii) of Theorem 2), then the government optimally censors all media outlets. In the intermediate case (part (ii) of Theorem 2), the government optimally permits only one moderate media outlet.

Roughly speaking, the case of unimodal V corresponds to a homogenous society where most citizens are moderates, and the case of bimodal V corresponds to a polarized society where most citizens are either extreme supporters or extreme opponents. It seems intuitive that the government optimally censors the least supportive media outlets, when the society is homogenous. But it may seem counterintuitive that the government may also optimally censor the most supportive media outlets, when the society is polarized. Such a censorship policy ensures that extreme supporters continue to choose the government's preferred action even if no permitted media outlets endorse it.

REFERENCES

ALONSO, RICARDO AND ODILON CÂMARA (2016): "Political Disagreement and Information in Elections," *Games and Economic Behavior*, 100, 390–412. [3]

ARIELI, ITAI, YAKOV BABICHENKO, RANN SMORODINSKY, AND TAKURO YAMASHITA (2023): "Optimal Persuasion via Bi-Pooling," *Theoretical Economics*, 18, 15–36. [2]

AYBAS, YUNUS C AND ERAY TURKEL (2024): "Persuasion with Coarse Communication," Working Paper. [1]

BLACKWELL, DAVID (1951): "Comparison of Experiments," in *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1, 93–102. [2]

DWORCZAK, PIOTR AND GIORGIO MARTINI (2019): "The Simple Economics of Optimal Persuasion," *Journal of Political Economy*, 127, 1993–2048. [1, 3]

GENTZKOW, MATTHEW AND EMIR KAMENICA (2016): "A Rothschild-Stiglitz Approach to Bayesian Persuasion," *American Economic Review, Papers & Proceedings*, 106, 597–601. [1, 12]

HOPENHAYN, HUGO AND MARYAM SAEEDI (2024): "Optimal Simple Ratings," Working Paper. [1]

IVANOV, MAXIM (2021): "Optimal Monotone Signals in Bayesian Persuasion Mechanisms," *Economic Theory*, 72 (3), 955–1000. [1]

KOLOTILIN, ANTON (2018): "Optimal Information Disclosure: A Linear Programming Approach," *Theoretical Economics*, 13, 607–636. [1, 2, 5]

KOLOTILIN, ANTON AND HONGYI LI (2021): "Relational Communication," *Theoretical Economics*, 16 (4), 1391–1430. [1]

KOLOTILIN, ANTON, TIMOFIY MYLOVANOV, AND ANDRIY ZAPECHELNYUK (2022): "Censorship as Optimal Persuasion," *Theoretical Economics*, 17, 561–585. [1, 2, 3, 6, 8]

KOLOTILIN, ANTON, TYMOFIY MYLOVANOV, ANDRIY ZAPECHELNYUK, AND MING LI (2017): "Persuasion of a Privately Informed Receiver," *Econometrica*, 85, 1949–1964. [1, 12]

KOLOTILIN, ANTON AND ANDRIY ZAPECHELNYUK (2024): "Persuasion Meets Delegation," Working Paper. [2]

LYU, QIANJUN, WING SUEN, AND YIMENG ZHANG (2024): "Coarse Information Design," Working Paper. [1]

MENSCH, JEFFREY (2021): "Monotone Persuasion," *Games and Economic Behavior*, 130, 521–542. [1]

ONUCHIC, PAULA AND DEBRAJ RAY (2023): "Conveying Value via Categories," *Theoretical Economics*, 18, 1407–1439. [1]

[4]Formally, a bimodal V can be concave-convex, concave-convex-concave, convex-concave-convex, or convex-concave-convex-concave. For simplicity, we restrict attention to the concave-convex-concave case, but our analysis can be extended to cover all cases.

APPENDIX A: PROOFS

A.1. *Single-Crossing Lemma*

We present a lemma adapted from Kolotilin et al. (2022), which is used in the proofs of Theorems 1 and 2.

Define $\Delta(\omega,m)=V(\omega)-V(m)-V'(m)(\omega-m)$.

LEMMA 1: *Let $0\leq\omega_L<\omega_R<1$. Suppose that V is strictly convex on $[\omega_L,\omega_R]$ and strictly concave on $[\omega_R,1]$. Then $\Delta(\omega,m)<0$ for all $\omega,m\in[\omega_R,1]$, with $\omega<m$. Moreover, if $\omega_L\leq\omega_1\leq\omega_2<\omega_R\leq m_1\leq m_2\leq 1$, with $(\omega_1,m_1)\neq(\omega_2,m_2)$, then*

$$\Delta(\omega_1,m_1)\leq 0 \implies \Delta(\omega_2,m_2)<0. \tag{7}$$

PROOF: By integration by parts,

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

Since V is strictly concave on $[\omega_R,1]$, we have $V''(z)<0$ for almost all $z\in[\omega_R,1]$, and hence $\Delta(\omega,m)<0$ for all $\omega,m\in[\omega_R,1]$, with $\omega<m$. Next,

$$\begin{aligned}\Delta(\omega_2,m_2)=\int_{\omega_2}^{m_2}V''(z)(z-\omega_2)\mathrm{d}z &\leq \int_{\omega_2}^{m_1}V''(z)(z-\omega_2)\mathrm{d}z\\ =\int_{\omega_2}^{m_1}V''(z)(z-\omega_1)\frac{z-\omega_2}{z-\omega_1}\mathrm{d}z &\leq \frac{\omega_R-\omega_2}{\omega_R-\omega_1}\int_{\omega_2}^{m_1}V''(z)(z-\omega_1)\mathrm{d}z\\ &\leq \frac{\omega_R-\omega_2}{\omega_R-\omega_1}\int_{\omega_1}^{m_1}V''(z)(z-\omega_1)\mathrm{d}z\leq 0,\end{aligned}$$

where the first inequality holds because $\omega_R\leq m_1\leq m_2\leq 1$ and V is concave on $[\omega_R,1]$, the second inequality holds because V is convex on $[\omega_L,\omega_R]$, concave on $[\omega_R,1]$, and $(z-\omega_2)/(z-\omega_1)$ is increasing in z, the third inequality holds because $\omega_L\leq\omega_1\leq\omega_2<\omega_R$ and V is convex on $[\omega_L,\omega_R]$, and the fourth inequality holds by $\Delta(\omega_1,m_1)\leq 0$. Moreover, $\Delta(\omega_2,m_2)<0$ if $\omega_1<\omega_2$ because the third inequality is strict by strict convexity of V on $[\omega_L,\omega_R]$, and $\Delta(\omega_2,m_2)<0$ if $m_1<m_2$ because the first inequality is strict by strict concavity of V on $[\omega_R,1]$. *Q.E.D.*

A.2. *Proof of Theorem 1*

When the state is discrete, the problem of finding an optimal monotone signal is a discrete optimization problem, which cannot be solved using the existing tools from the persuasion literature. To prove Theorem 1, we first show that any monotone signal different from deterministic upper censorship is suboptimal, and then we show that optimal deterministic upper censorship has the same cutoff as the stochastic upper censorship signal that solves the unrestricted persuasion problem.

PROOF OF THEOREM 1: Let $\text{supp}(F) = \{\omega_1, \ldots, \omega_n\}$, with natural n and $\omega_1 < \ldots < \omega_n$. For each $1 \leq i < j \leq n$, denote $f_j = f(\omega_j)$, $f_{i:j} = \mathbb{P}(\omega \in \{\omega_i, \ldots, \omega_j\})$, and $m_{i:j} = \mathbb{E}[\omega | \omega \in \{\omega_i, \ldots, \omega_j\}]$.

Suppose by contradiction that there exists a monotone signal μ that is not deterministic upper censorship. Then there exist $1 \leq i < j < k \leq n$ and two signal realizations: s_1 that pools states $\{\omega_i, \ldots, \omega_j\}$ and s_2 that pools states $\{\omega_{j+1}, \ldots, \omega_k\}$. Let μ_- and μ_+ be monotone signals that differ from μ only in that μ_- merges signal realizations s_1 and s_2 of μ into one signal realization that pools states $\{\omega_i, \ldots, \omega_k\}$ and μ_+ splits signal realization s_1 of μ into two signal realizations: one that separates state ω_j and one that pools remaining states $\{\omega_i, \ldots, \omega_{j-1}\}$. Denote the value of signals μ_-, μ, and μ_+ by W_-, W, and W_+.

To obtain a contradiction, it suffices to show that $W \geq W_+$ implies $W < W_-$. So, suppose that $W \geq W_+$, which is equivalent to

$$V(m_{i:j}) \geq \frac{f_{i:j-1}}{f_{i:j}} V(m_{i:j-1}) + \frac{f_j}{f_{i:j}} V(\omega_j). \tag{8}$$

Since V is strictly convex on $[0, \omega_M]$ and (8) holds, it follows that $\omega_M < \omega_j$.

We now show that $W < W_-$, which is equivalent to

$$V(m_{i:k}) > \frac{f_{i:j}}{f_{i:k}} V(m_{i:j}) + \frac{f_{j+1:k}}{f_{i:k}} V(m_{j+1:k}). \tag{9}$$

If $\omega_M \leq m_{i:j}$, then (9) follows from strict concavity of V on $[\omega_M, 1]$. So, suppose that $\omega_M \in (m_{i:j}, \omega_j)$. By $\omega_M < \omega_j < m_{j+1:k}$ and strict concavity of V on $[\omega_M, 1]$, we have

$$V(\omega_M) < \frac{m_{j+1:k} - \omega_M}{m_{j+1:k} - \omega_j} V(\omega_j) + \frac{\omega_M - \omega_j}{m_{j+1:k} - \omega_j} V(m_{j+1:k}). \tag{10}$$

We also have

$$\begin{aligned} V(\omega_M) &> \frac{m_{i:j} - \omega_M}{m_{i:j} - m_{i:j-1}} V(m_{i:j-1}) + \frac{\omega_M - m_{i:j-1}}{m_{i:j} - m_{i:j-1}} V(m_{i:j}) \\ &\geq \frac{\omega_j - \omega_M}{\omega_j - m_{i:j-1}} V(m_{i:j-1}) + \frac{\omega_M - m_{i:j-1}}{\omega_j - m_{i:j-1}} V(\omega_j), \end{aligned} \tag{11}$$

where the first inequality is by $m_{i:j-1} < m_{i:j} < \omega_M$ and strictly convexity of V on $[0, \omega_M]$, and the second inequality is by (8). Combining (10) and (11) yields

$$V(\omega_M) > \frac{m_{j+1:k} - \omega_M}{m_{j+1:k} - m_{i:j-1}} V(m_{i:j-1}) + \frac{\omega_M - m_{i:j-1}}{m_{j+1:k} - m_{i:j-1}} V(m_{j+1:k}). \tag{12}$$

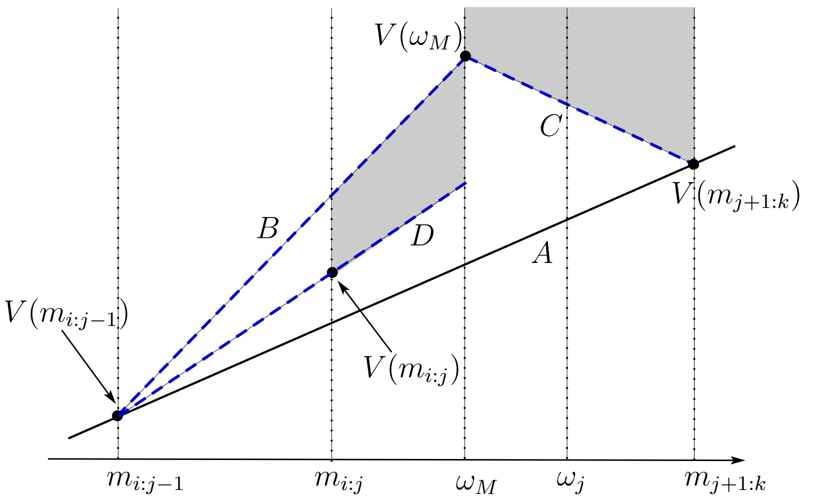


FIGURE A.1.

As illustrated in Figure A.1, by (12), $(\omega_M, V(\omega_M))$ is above the solid line A connecting $(m_{i:j-1}, V(m_{i:j-1}))$ and $(m_{j+1:k}, V(m_{j+1:k}))$. Then, by strict concavity of V on $[\omega_M, 1]$, $(\omega_j, V(\omega_j))$ is above the dashed line C connecting $(\omega_M, V(\omega_M))$ and $(m_{j+1:k}, V(m_{j+1:k}))$, and thus above the line A. So, by (8), $(m_{i:j}, V(m_{i:j}))$ is also above the line A. Next, observe that $(m_{i:k}, V(m_{i:k}))$ must be in the shaded area in Figure A.1. Indeed, if $m_{i:k} < \omega_M$, then, by strict convexity of V on $[0, \omega_M]$, $(m_{i:k}, V(m_{i:k}))$ is above the dashed line D connecting $(m_{i:j-1}, V(m_{i:j-1}))$ and $(m_{i:j}, V(m_{i:j}))$. If $m_{i:k} > \omega_M$, then, by strict concavity of V on $[\omega_M, 1]$, $(m_{i:k}, V(m_{i:k}))$ is above the line C. So, (9) holds.

We now show that the optimal deterministic upper censorship cutoff coincides with the optimal stochastic upper censorship cutoff. For each $z \in [\omega_1, \omega_n]$, define

$$j(z) = \max\{i \in \{1, \dots, n\} : \omega_i \le z\},$$

$$q(z) = \frac{z - \omega_{j(z)}}{\omega_{j(z)+1} - \omega_{j(z)}},$$

$$m(z) = \frac{(1 - q(z)) f_{j(z)} \omega_{j(z)} + \sum_{i > j(z)} f_i \omega_i}{(1 - q(z)) f_{j(z)} + \sum_{i > j(z)} f_i},$$

$$W(z) = \sum_{i < j(z)} f_i V(\omega_i) + q(z) f_{j(z)} V(\omega_{j(z)}) + \left((1 - q(z)) f_{j(z)} + \sum_{i > j(z)} f_i \right) V(m(z)).$$

Thus, every $z \in [\omega_1, \omega_n]$ represents a stochastic upper censorship signal with $(\omega_{j(z)}, q(z))$, where $m(z)$ is the expected state conditional on the pooling signal realization and $W(z)$ is the value of this signal. Conversely, every stochastic upper censorship signal with (ω_j, q) in

$\{\omega_1,...,\omega_{n-1}\}\times[0,1]$ can be represented by $z=(1-q)\omega_j+q\omega_{j+1}\in[\omega_1,\omega_n]$.[5] Also note that z represents deterministic upper censorship iff $z\in\{\omega_1,...,\omega_n\}$.

Observe that $m(z)$ and $W(z)$ are continuous by construction. Taking the derivative at $z\notin\{\omega_1,\dots,\omega_n\}$, we obtain

$$m'(z)=\frac{q'(z)\sum_{i>j(z)}f_i(\omega_i-\omega_{j(z)})}{\left((1-q(z))f_{j(z)}+\sum_{i>j(z)}f_i\right)^2}=\frac{q'(z)f_{j(z)}(m(z)-\omega_{j(z)})}{(1-q(z))f_{j(z)}+\sum_{i>j(z)}f_i},$$

$$\begin{aligned}W'(z)&=q'(z)f_{j(z)}(V(\omega_{j(z)})-V(m(z))+m'(z)\left((1-q(z))f_{j(z)}+\sum_{i>j(z)}f_i\right)V'(m(z))\\&=q'(z)f_{j(z)}\big(V(\omega_{j(z)})-V(m(z))-V'(m(z))(\omega_{j(z)}-m(z))\big)\\&=\frac{f_{j(z)}}{\omega_{j(z)+1}-\omega_{j(z)}}\Delta(\omega_{j(z)},m(z)).\end{aligned}$$

Since $\omega_{j(z)}$ and $m(z)$ are increasing in z and $f_{j(z)}/(\omega_{j(z)+1}-\omega_{j(z)})>0$, Lemma 1 implies that $W'(z)$ is strictly downcrossing in z. This implies that the optimal unrestricted signal is unique and is stochastic upper censorship with some cutoff ω^*. Furthermore, this implies that an optimal monotone signal is deterministic upper censorship with the same cutoff ω^* and some $q^{**}\in\{0,1\}$. *Q.E.D.*

A.3. *Proof of Theorem 2*

When the monotonicity constraint is binding (Figures 1c and 1d), the existing approaches from the Bayesian persuasion literature no longer apply to the monotone persuasion problem. We first consider a constrained monotone persuasion problem with the two additional constraints that monotone signals must satisfy at inflection points ω_L and ω_R. We show that an optimal solution to this constrained problem takes the form of interval disclosure or partitions (almost) all the states into three pooling intervals. Finally, to obtain parts (ii) and (iii) of Theorem 2, we consider the original monotone persuasion problem without additional constraints, and show that a monotone signal consisting of three pooling intervals is suboptimal.

It is convenient to represent a monotone signal by a *pooling set* $P\subset[0,1]$ of states that are not separated by this signal. Since the state is continuous, w.l.o.g., each pooling interval is open. Thus, the pooling set is a union of some disjoint nonempty open intervals, $P=\bigcup_i(\xi_i,\zeta_i)$. Let $\mathcal{P}_{[0,1]}$ be the set of all open subsets of $[0,1]$.[6] A pooling set $P\in\mathcal{P}_{[0,1]}$ corresponds to the monotone signal μ_P given by

$$\mu_P(\omega)=\begin{cases}\omega, & \omega\notin(\xi_i,\zeta_i)\text{ for all }i,\\ \mathbb{E}[\omega|(\xi_i,\zeta_i)], & \omega\in(\xi_i,\zeta_i)\text{ for some }i.\end{cases}$$

[5] An upper censorship with (ω_n,q), for any $q\in[0,1]$, is the same as the upper censorship with $(\omega_{n-1},1)$.

[6] We define open sets in $[0,1]$ rather than in $\mathbb{R}$; e.g., $[0,1/2)\cup(1/2,1]$ is open.

The distribution G_P of $\mu_P(\omega)$ is given by

$$G_P(\omega) = \begin{cases} F(\omega), & \text{if } \omega \notin (\xi_i, \zeta_i) \text{ for all } i, \\ F(\xi_i), & \text{if } \omega \in (\xi_i, \mathbb{E}[\omega|(\xi_i, \zeta_i)]) \text{ for some } i, \\ F(\zeta_i), & \text{if } \omega \in [\mathbb{E}[\omega|(\xi_i, \zeta_i)], \zeta_i) \text{ for some } i. \end{cases}$$

Solving the monotone persuasion problem is thus equivalent to finding an *optimal* pooling set P^* that maximizes $\int_0^1 V(\omega)\mathrm{d}G_P(\omega)$ over $P \in \mathcal{P}_{[0,1]}$.

Following Gentzkow and Kamenica (2016) and Kolotilin et al. (2017), we represent signals as convex functions. Given a pooling set P, define

$$\Gamma_P(\omega) = \int_0^\omega G_P(\tilde{\omega})\mathrm{d}\tilde{\omega} \quad \text{for all } \omega \in [0,1].$$

Similarly to Kolotilin et al. (2017), we have the following optimality criterion.

LEMMA 2: *A pooling set P^* is optimal iff it solves*

$$\max_{P \in \mathcal{P}_{[0,1]}} \int_0^1 V''(\omega)\Gamma_P(\omega)\mathrm{d}\omega. \tag{13}$$

PROOF: Since V is twice continuously differentiable, we can integrate by parts twice,

$$\begin{aligned} \int_0^1 V(\omega)\mathrm{d}G_P(\omega) &= V(\omega)G_P(\omega)|_0^1 - \int_0^1 V'(\omega)G_P(\omega)\mathrm{d}\omega \\ &= V(\omega)G_P(\omega)|_0^1 - V'(\omega)\Gamma_P(\omega)|_0^1 + \int_0^1 V''(\omega)\Gamma_P(\omega)\mathrm{d}\omega \\ &= V(1) - V'(1)(1-\mathbb{E}[\omega]) + \int_0^1 V''(\omega)\Gamma_P(\omega)\mathrm{d}\omega, \end{aligned} \tag{14}$$

where the last equality follows from

$$\Gamma_P(1) = \int_0^1 G_P(\omega)\mathrm{d}\omega = \omega G_P(\omega)|_0^1 - \int_0^1 \omega \mathrm{d}G_P(\omega) = 1 - \mathbb{E}[\omega].$$

Since only the last term of (14) depends on P, the proposition follows. *Q.E.D.*

As (13) suggests, the optimal pooling set P^* should be chosen to make $\Gamma_P(\omega)$ large at states ω where $V''(\omega)$ is positive, and small at states where $V''(\omega)$ is negative. Separating state ω increases $\Gamma_P(\omega)$, so full disclosure ($P^* = \varnothing$) is optimal iff $V(\omega)$ is convex in ω. In contrast, no disclosure ($P^* = [0,1]$) is optimal if $V(\omega)$ is concave in ω. These conditions for the optimality of full disclosure and no disclosure follow easily from Lemma 2. In turn, Lemma 3 collects useful properties of the function Γ_P.

LEMMA 3: *For all $P \in \mathcal{P}_{[0,1]}$,*

(i) $\Gamma_P(\omega)$ *is convex in* ω.
(ii) $\Gamma_{[0,1]}(\omega) \leq \Gamma_P(\omega) \leq \Gamma_\emptyset(\omega)$ *for all* $\omega \in [0,1]$.
(iii) $\Gamma_P(\omega) = \Gamma_\emptyset(\omega)$ *iff* $\omega \notin P$.

PROOF: Part (i) holds because $\Gamma_P(\omega) = \int_0^\omega G_P(\tilde{\omega})\mathrm{d}\tilde{\omega}$ and $G_P(\omega)$ is a (non-decreasing) distribution function. For parts (ii) and (iii), we first show that

$$\int_0^\omega G_P(\tilde{\omega})\mathrm{d}\tilde{\omega} = \Gamma_P(\omega) \leq \Gamma_\emptyset(\omega) = \int_0^\omega F(\tilde{\omega})\mathrm{d}\tilde{\omega} \text{ for all } \omega \in [0,1],$$

with equality iff $\omega \notin P$. It is sufficient to observe that for each disjoint interval (ξ_i, ζ_i) of P, we have

$$\int_{\xi_i}^\omega G_P(\tilde{\omega})\mathrm{d}\tilde{\omega} = F(\xi_i)(\omega - \xi_i) < \int_{\xi_i}^\omega F(\tilde{\omega})\mathrm{d}\tilde{\omega} \text{ for } \omega \in (\xi_i, \mathbb{E}[\omega|(\xi_i,\zeta_i)]),$$
$$\int_\omega^{\zeta_i} G_P(\tilde{\omega})\mathrm{d}\tilde{\omega} = F(\zeta_i)(\zeta_i - \omega) > \int_\omega^{\zeta_i} F(\tilde{\omega})\mathrm{d}\tilde{\omega} \text{ for } \omega \in [\mathbb{E}[\omega|(\xi_i,\zeta_i)], \zeta_i),$$
$$\int_{\xi_i}^{\zeta_i} G_P(\tilde{\omega})\mathrm{d}\tilde{\omega} = F(\xi_i)(\mathbb{E}[\omega|(\xi_i,\zeta_i)] - \xi_i) + F(\zeta_i)(\zeta_i - \mathbb{E}[\omega|(\xi_i,\zeta_i)]) = \int_{\xi_i}^{\zeta_i} F(\tilde{\omega})\mathrm{d}\tilde{\omega},$$

where each line holds, respectively, because

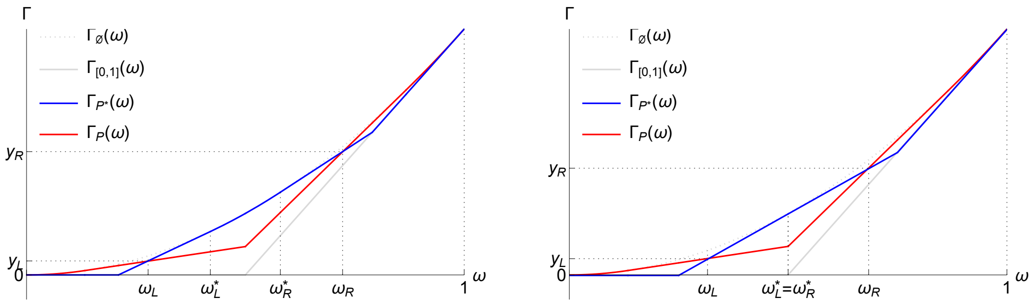
$$F(\xi_i) < F(\omega) \text{ for } \omega \in (\xi_i, \mathbb{E}[\omega|(\xi_i,\zeta_i)]),$$
$$F(\zeta_i) > F(\omega) \text{ for } \omega \in [\mathbb{E}[\omega|(\xi_i,\zeta_i)], \zeta_i),$$
$$\int_{\xi_i}^{\zeta_i} F(\tilde{\omega})\mathrm{d}\tilde{\omega} = F(\omega)\omega|_{\xi_i}^{\zeta_i} - \int_{\xi_i}^{\zeta_i} \tilde{\omega}\mathrm{d}F(\tilde{\omega}) = F(\zeta_i)\zeta_i - F(\xi_i)\xi_i - (F(\zeta_i) - F(\xi_i))\mathbb{E}[\omega|(\xi_i,\zeta_i)].$$

Similarly, the remainder of part (ii) that $\Gamma_{[0,1]}(\omega) \leq \Gamma_P(\omega)$ for all $\omega \in [0,1]$ follows from

$$\int_0^\omega G_{[0,1]}(\tilde{\omega})\mathrm{d}\tilde{\omega} \leq \int_0^\omega G_P(\tilde{\omega})\mathrm{d}\tilde{\omega} \text{ for } \omega \in (0, \mathbb{E}[\omega]),$$
$$\int_\omega^1 G_{[0,1]}(\tilde{\omega})\mathrm{d}\tilde{\omega} \geq \int_\omega^1 G_P(\tilde{\omega})\mathrm{d}\tilde{\omega} \text{ for } \omega \in [\mathbb{E}[\omega], 1),$$
$$\int_0^1 G_{[0,1]}(\tilde{\omega})\mathrm{d}\tilde{\omega} = \int_0^1 G_P(\tilde{\omega})\mathrm{d}\tilde{\omega},$$

where each line holds, respectively, because

$$G_{[0,1]}(\omega) = 0 \leq G_P(\omega) \text{ for } \omega \in (0, \mathbb{E}[\omega]),$$
$$G_{[0,1]}(\omega) = 1 \geq G_P(\omega) \text{ for } \omega \in [\mathbb{E}[\omega], 1),$$



(a) $\omega_L^* < \omega_R^*$ (b) $\omega_L^* = \omega_R^*$

FIGURE A.2.—Optimal pooling set P^* given $\omega_L^* \leq \omega_R^*$

$$\int_0^1 G_P(\omega)\mathrm{d}\omega = \omega G_P(\omega)|_0^1 - \int_0^1 \omega \mathrm{d}G_P(\omega) = 1 - \mathbb{E}[\omega]. \qquad \textit{Q.E.D.}$$

PROOF OF THEOREM 2: Define Y as the set of pairs $(y_L, y_R) \in \mathbb{R}_+^2$ such that $\Gamma_P(\omega_L) = y_L$ and $\Gamma_P(\omega_R) = y_R$ for some $P \in \mathcal{P}_{[0,1]}$. Fix $(y_L, y_R) \in Y$. We first consider problem (13) subject to the two additional constraints that $\Gamma_P(\omega_L) = y_L$ and $\Gamma_P(\omega_R) = y_R$.

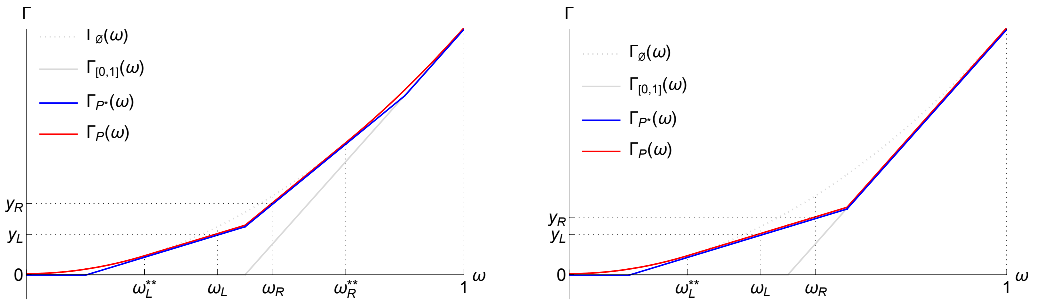
$$P^* \in \arg\max_{P \in \mathcal{P}_{[0,1]}} \int_0^1 V''(\omega)\Gamma_P(\omega)d\omega \quad \text{subject to } \Gamma_P(\omega_L) = y_L \text{ and } \Gamma_P(\omega_R) = y_R. \tag{15}$$

Let $\omega_L^* \geq \omega_L$ be the tangency point of a tangent line to $\Gamma_\emptyset$ that passes through the point (ω_L, y_L), and let $\omega_R^* \leq \omega_R$ be the tangency point of a tangent line to $\Gamma_\emptyset$ that passes through the point (ω_R, y_R) (Figure A.2). Formally, define

$$\omega_L^* = \min\{\omega \in [\omega_L, 1] : \Gamma_{[0,\omega)}(\omega_L) = y_L\},$$
$$\omega_R^* = \max\{\omega \in [0, \omega_R] : \Gamma_{(\omega,1]}(\omega_R) = y_R\}.$$

CLAIM 1: *If $\omega_L^* \leq \omega_R^*$, then $P^* = [0, \omega_L^*) \cup (\omega_R^*, 1]$ solves (15).*

PROOF: By Lemma 3, for any $P \in \mathcal{P}_{[0,1]}$ such that $\Gamma_P(\omega_L) = y_L$ and $\Gamma_P(\omega_R) = y_R$, we have $\Gamma_P(\omega)$ is convex in ω and $\Gamma_{[0,1]}(\omega) \leq \Gamma_P(\omega) \leq \Gamma_\emptyset(\omega)$ for all $\omega \in [0,1]$. It is easy to verify (Figure A.2) that for any such Γ_P, we have $\Gamma_{P^*}(\omega) \leq \Gamma_P(\omega)$ for $\omega \in [0, \omega_L) \cup (\omega_R, 1]$ and $\Gamma_{P^*}(\omega) \geq \Gamma_P(\omega)$ for $\omega \in (\omega_L, \omega_R)$. Since $V''(\omega) \leq 0$ for $\omega \in [0, \omega_L) \cup (\omega_R, 1]$ and $V''(\omega) \geq 0$ for $\omega \in (\omega_L, \omega_R)$, the set P^* solves (15). *Q.E.D.*



(a) $\Gamma_{(\omega_L^{**},1]}(\omega_R) \neq y_R$ and $\Gamma_{[0,\omega_R^{**})}(\omega_L) \neq y_L$ (b) $\Gamma_{(\omega_L^{**},1]}(\omega_R) = y_R$

FIGURE A.3.—Optimal pooling set P^* given $\omega_L^* > \omega_R^*$

Let $\omega_L^{**} \leq \omega_L$ be the tangency point of a tangent line to $\Gamma_\emptyset$ that passes through the point (ω_L, y_L), and let $\omega_R^{**} \geq \omega_R$ be the tangency point of a tangent line to $\Gamma_\emptyset$ that passes through the point (ω_R, y_R) (Figure A.3a). Formally, define

$$\begin{aligned} \omega_L^{**} &= \min\{\omega \in [0, \omega_L] : \Gamma_{(\omega,1]}(\omega_L) = y_L\}, \\ \omega_R^{**} &= \max\{\omega \in [\omega_R, 1] : \Gamma_{[0,\omega)}(\omega_R) = y_R\}. \end{aligned} \tag{16}$$

CLAIM 2: *Suppose* $\omega_L^* > \omega_R^*$.
(i) *If* $\Gamma_{(\omega_L^{**},1]}(\omega_R) = y_R$, *then* $P^* = [0, \omega_L^{**}) \cup (\omega_L^{**}, 1]$ *solves (15).*
(ii) *If* $\Gamma_{[0,\omega_R^{**})}(\omega_L) = y_L$, *then* $P^* = [0, \omega_R^{**}) \cup (\omega_R^{**}, 1]$ *solves (15).*
(iii) *Otherwise,* $P^* = [0, \omega_L^{**}) \cup (\omega_L^{**}, \omega_R^{**}) \cup (\omega_R^{**}, 1]$ *solves (15).*

PROOF: The proof of parts (i) and (ii) is analogous to the proof of Claim 1 (Figure A.3b).

We now outline the proof of part (iii), omitting tedious details. The reader may refer to Figure A.3a for guidance. If $\omega_L^* > \omega_R^*$ with $(y_L, y_R) \in Y$, then

$$y_L + \frac{y_R - y_L}{\omega_R - \omega_L}(\omega - \omega_L) < \Gamma_\emptyset(\omega) \text{ for } \omega \in [\omega_L, \omega_R]. \tag{17}$$

Taking into account (17), if $\Gamma_{(\omega_L^{**},1]}(\omega_R) \neq y_R$ and $\Gamma_{[0,\omega_R^{**})}(\omega_L) \neq y_L$ with $(y_L, y_R) \in Y$, then $\omega_L^{**} \in [0, \omega_L)$ and $\omega_R^{**} \in (\omega_R, 1]$. We can then show, using the definitions of G_P and Γ_P, that $\Gamma_P(\omega_L) = y_L$ and $\Gamma_P(\omega_R) = y_R$ with $(y_L, y_R) \in Y$ iff $(\omega_L^{**}, \omega_R^{**})$ is a disjoint interval in P. By Lemma 3, for any such P, we have $\Gamma_{P^*}(\omega) = \Gamma_P(\omega)$ for $\omega \in [\omega_L^{**}, \omega_R^{**}]$ and $\Gamma_{P^*}(\omega) \leq \Gamma_P(\omega)$ for $\omega \in [0, \omega_L^{**}) \cup (\omega_R^{**}, 1]$. Since $V''(\omega) \leq 0$ for $\omega \in [0, \omega_L^{**}) \cup (\omega_R^{**}, 1] \subset [0, \omega_L) \cup (\omega_R, 1]$, the set P^* solves (15). *Q.E.D.*

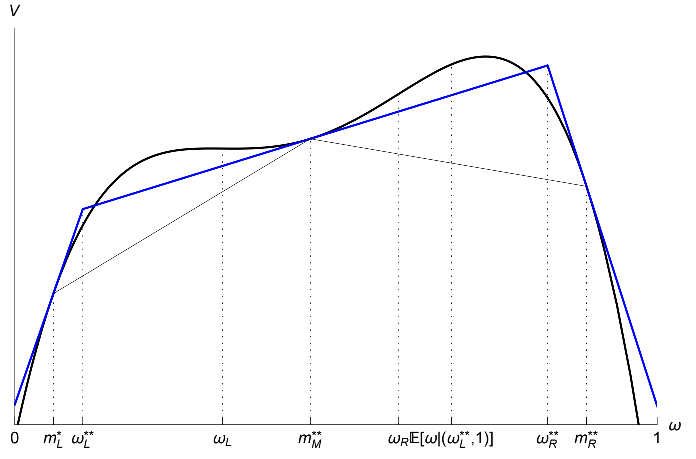


FIGURE A.4.—Optimal pooling set $P^{**} = [0, \omega_L^{**}) \cup (\omega_L^{**}, \omega_R^{**}) \cup (\omega_R^{**}, 1]$

We now consider the original problem (13), without the constraints that $\Gamma_P(\omega_L) = y_L$ and $\Gamma_P(\omega_R) = y_R$.

CLAIM 3: *Suppose* $P^{**} = [0, \omega_L^{**}) \cup (\omega_L^{**}, \omega_R^{**}) \cup (\omega_R^{**}, 1]$ *solves* (13) *where* $\omega_L^{**} > 0$, $\omega_R^{**} < 1$, *and (16) holds with* $y_L = \Gamma_{P^*}(\omega_L)$ *and* $y_R = \Gamma_{P^*}(\omega_R)$. *Let* $m_M^{**} = \mathbb{E}[\omega|(\omega_L^{**}, \omega_R^{**})]$. *If* $m_M^{**} > \omega_L$, *then* $P^* = [0, \omega_L^{**}) \cup (\omega_L^{**}, 1]$ *also solves* (13). *If* $m_M^{**} < \omega_R$, *then* $P^* = [0, \omega_R^{**}) \cup (\omega_R^{**}, 1]$ *also solves* (13).

PROOF: By (13), we have $\omega_L^{**} \leq \omega_L$ and $\omega_R^{**} \geq \omega_R$. Thus, $m_L^{**} < \omega_L$ and $m_R^{**} > \omega_R$, where $m_L^{**} = \mathbb{E}[\omega|[0, \omega_L^{**})]$ and $m_R^{**} = \mathbb{E}[\omega|(\omega_R^{**}, 1]]$.

The value of $P^{**} = [0, \omega_L^{**}) \cup (\omega_L^{**}, \omega_R^{**}) \cup (\omega_R^{**}, 1]$ is

$$
\begin{aligned}
v^{**} &= \int_0^1 V(\omega) \mathrm{d}G_{P^{**}}(\omega) \\
&= V(m_L^{**}) F(\omega_L^{**}) + V(m_M^{**}) (F(\omega_R^{**}) - F(\omega_L^{**})) + V(m_R^{**}) (1 - F(\omega_R^{**})).
\end{aligned}
$$

Since ω_R^{**} is interior, it satisfies the following first-order condition. Taking into account that

$$
\frac{\mathrm{d}m_M^{**}}{\mathrm{d}\omega_R^{**}} = \frac{f(\omega_R^{**})}{F(\omega_R^{**}) - F(\omega_L^{**})} (\omega_R^{**} - m_M^{**}) \quad \text{and} \quad \frac{\mathrm{d}m_R^{**}}{\mathrm{d}\omega_R^{**}} = \frac{f(\omega_R^{**})}{1 - F(\omega_R^{**})} (m_R^{**} - \omega_R^{**}),
$$

we have

$$
\frac{\mathrm{d}v^{**}}{\mathrm{d}\omega_R^{**}} = f(\omega_R^{**}) \left(V(m_M^{**}) + V'(m_M^{**})(\omega_R^{**} - m_M^{**}) - V(m_R^{**}) - V'(m_R^{**})(\omega_R^{**} - m_R^{**}) \right) = 0,
$$

which can be rewritten as

$$V(m_M^{**}) + V'(m_M^{**})(\omega_R^{**} - m_M^{**}) = V(m_R^{**}) + V'(m_R^{**})(\omega_R^{**} - m_R^{**}). \tag{18}$$

Consider now the case where $m_M^{**} > \omega_L$. Combining (18) with the fact that V is either concave on $(m_M^{**}, 1]$ or convex on (m_M^{**}, ω_R) and concave on $(\omega_R, 1]$, Lemma 1 (Figure A.4) implies that

$$V(\omega) \geq \frac{m_R^{**} - \omega}{m_R^{**} - m_M^{**}} V(m_M^{**}) + \frac{\omega - m_M^{**}}{m_R^{**} - m_M^{**}} V(m_R^{**}) \text{ for } \omega \in (m_M^{**}, m_R^{**}),$$

which, given $\mathbb{E}[\omega|(\omega_L^{**}, 1]] \in (m_M^{**}, m_R^{**})$, implies, for $P^* = [0, \omega_L^{**}) \cup (\omega_L^{**}, 1]$,

$$\begin{aligned}\int_0^1 V(\omega)\mathrm{d}G_{P^*}(\omega) &= V(m_L^{**})F(\omega_L^{**}) + V(\mathbb{E}[\omega|(\omega_L^{**}, 1]])(1 - F(\omega_L^{**})) \\ &\geq V(m_L^{**})F(\omega_L^{**}) + V(m_\star^{M})(F(\omega_R^{**}) - F(\omega_L^{**})) + V(m_R^{**})(1 - F(\omega_R^{**})) \\ &= \int_0^1 V(\omega)\mathrm{d}G_{P^{**}}(\omega),\end{aligned}$$

showing that P^* also solves (13), if $m_M^{**} > \omega_L$. A symmetric argument applies when $m_M^{**} < \omega_R$. *Q.E.D.*

Combining Claims 1 – 3, we conclude that P^* takes one of three forms: $[0, \omega_L^*) \cup (\omega_R^*, 1]$ with $\omega_L \leq \omega_L^* < \omega_R^* \leq \omega_R$, or $[0, \omega_M^*) \cup (\omega_M^*, 1]$ with $\omega_M^* \in (0, 1)$, or $[0, 1]$. *Q.E.D.*

A.4. *Proof of Proposition 1*

Let $X \subset C$. For each $\omega \in [0,1]$, define $\underline{c}_X(\omega) = \sup(\{c \in C \backslash X : c \leq \omega\} \cup \{0\})$, $\overline{c}_X(\omega) = \inf(\{c \in C \backslash X : c > \omega\} \cup \{1\})$, and $\mu(\omega) = \mathbb{E}[\theta|\theta \in [\underline{c}_X(\omega), \overline{c}_X(\omega)]]$. Observe that μ is a monotone signal such that $G_\mu = G_X$. Thus, $\mathcal{G}_C \subset \mathcal{G}_M$.

Conversely, let $\mu \in \mathcal{M}$. For each m such that $\mu(\omega) = m$ for some $\omega \in \mathrm{supp}(F)$, define $\underline{x}_\mu(m) = \inf\{\omega \in \mathrm{supp}(F) : \mu(\omega) = m\}$, $\overline{x}_\mu(m) = \sup\{\omega \in \mathrm{supp}(F) : \mu(\omega) = m\}$, and $X = \left(\bigcup_{m \in \mu(\mathrm{supp}(F))} (\underline{x}_\mu(m), \overline{x}_\mu(m)]\right) \bigcap C$. Observe that X is a censorship policy such that $G_X = G_\mu$. Thus, $\mathcal{G}_M \subset \mathcal{G}_C$. *Q.E.D.*