

ASSORTATIVE INFORMATION DISCLOSURE

ANTON KOLOTILIN AND ALEXANDER WOLITZKY

ABSTRACT. We consider a standard persuasion problem in which the receiver’s action and the state of the world are both one-dimensional. Fully characterizing optimal signals when utilities are non-linear is a daunting task. Instead, we develop a general approach to understanding a key qualitative property of optimal signals: their *assortative structure*, which describes the overall pattern of what states are pooled together. We derive simple conditions—driven by intuitive economic properties, such as supermodularity and submodularity of preferences—for the optimality of positive and negative assortative patterns of information disclosure. Our approach unifies a wide range of previous findings and generates new applications.

JEL Classification: C78, D82, D83

Keywords: persuasion, information design, assortative matching

Date: 21st November 2019.

Kolotilin: School of Economics, UNSW Business School.

Wolitzky: Department of Economics, MIT.

We thank Jaksa Cvitanik, Stephen Morris, and Eran Shmaya for helpful comments and suggestions. We thank Daniel Clark and Yucheng Shang for excellent research assistance. Anton Kolotilin gratefully acknowledges support from the Australian Research Council Discovery Early Career Research Award DE160100964 and from MIT Sloan’s Program on Innovation in Markets and Organizations. Alexander Wolitzky gratefully acknowledges support from NSF CAREER Award 1555071 and Sloan Foundation Fellowship 2017-9633.

1. INTRODUCTION

Following the seminal papers of [Rayo and Segal \(2010\)](#) and [Kamenica and Gentzkow \(2011\)](#), the past decade has witnessed an explosion of interest in the design of optimal information disclosure policies. Unfortunately, fully characterizing the optimal disclosure policy is usually a difficult problem. While significant progress has been made in the special case where the sender’s and receiver’s utilities are linear in the unknown state ([Gentzkow and Kamenica \(2016\)](#), [Kolotilin, Mylovanov, Zapechelnyuk, and Li \(2017\)](#), [Kolotilin \(2018\)](#), [Dworczak and Martini \(2019\)](#)), results with non-linear preferences are scarce, and it is not clear which insights from the linear case can be expected to hold more generally.

In this paper, rather than concerning ourselves with completely characterizing optimal signals, we develop a general approach to understanding a key qualitative property of such signals: their *assortative structure*, which describes the overall pattern of what states are pooled together, and how the induced receiver action varies over pairs of states. We obtain simple results driven by intuitive economic properties, such as supermodularity and submodularity of preferences.

We consider a standard persuasion problem with one sender (she) and one receiver (he). The receiver’s action and the state of the world are both one-dimensional. We impose two substantive assumptions. First, the sender’s utility is increasing in the receiver’s action. Second, the receiver’s optimal action is increasing in his belief about the state: more precisely, he prefers higher actions at higher states, and his expected utility is single-peaked in his action for any posterior belief.

We show that it is without loss to focus on *pairwise signals*, where each induced posterior distribution has at most binary support. That is, there always exists an optimal signal that pools at most two states on each signal realization, and typically every optimal signal takes this form. This conclusion differs starkly from that in the linear-utilities case, where it is often optimal to pool intervals of states (e.g., by “censoring” states on one side of a cutoff). Our results reveal that the optimality of interval pooling in this case rests on the “non-genericity” of linear utilities.¹

Restricting attention to pairwise signals reduces optimal persuasion to a generalized matching problem—what pairs of states should be matched together, and what probability should be assigned to each state in every pair? Our focus is characterizing conditions for three simple matching patterns to be optimal.

Positive assortative matching—where like states are matched together—corresponds to *full disclosure* of the state. We provide conditions for full disclosure to be optimal, which are closely related to conditions for optimality of positive assortativity in the

¹Of course, “not generic” does not mean “not important.” For example, the standard quasilinear utility assumption in mechanism design is “non-generic” in much the same way as linear utilities are non-generic in our model.

matching literature (Becker (1973), Legros and Newman (2002)). Our conditions generalize, for example, those in Rayo and Segal (2010) and Kolotilin (2018).

Our main results characterize *single-dipped disclosure* and *single-peaked disclosure*, which are forms of negative assortative disclosure where the supports of any two posterior distributions do not cross, so all pairs of supports are either ordered or nested. The difference between single-dipped and single-peaked disclosure is that, in the former case, the induced receiver action is single-dipped on each set of nested pairs of states; while in the latter case it is single-peaked on each such set. Our analysis of single-dipped and single-peaked disclosure unifies and generalizes most of the main qualitative results in the literature on persuasion with non-linear preferences, and also generates new results for previously unexplored persuasion models.

For instance, an example of single-dipped disclosure is Friedman and Holden’s (2008) “matching extremes” gerrymandering solution, where a gerrymanderer creates electoral districts that pool extreme supporters with similarly extreme opponents, and wins those districts with the most extreme supporters and opponents with the highest probability. A second example in the literature is Goldstein and Leitner’s (2018) non-monotone stress tests, where a regulator designs a stress test that pools the weakest banks that it wants to receive funding with the strongest banks, pools slightly less weak banks with slightly less strong banks, and so on, such that the weakest and strongest banks receive the highest funding. A novel application concerns a fund manager with access to an off-market investment opportunity that can profitably absorb a certain amount of capital, which is privately known to the manager. The manager’s optimal disclosure policy pools low and high-capacity opportunities, so that investors always invest just enough to fully exploit the high-capacity opportunity, but also overinvest with probability sufficient to drive their ex ante expected excess returns to zero. This disclosure policy is single-peaked, because investment is higher on signals that pool more extreme low and high-capacity opportunities.

Single-peaked disclosure also arises in applications. The main example we are aware of is Guo and Shmaya’s (2019) “nested intervals” disclosure rule, where an information designer creates signals that pool favorable states with similarly unfavorable states, and persuades the receiver to take her preferred action with the highest probability after those signals with the most moderate favorable and unfavorable states.

While our general characterization of single-dipped and single-peaked disclosure is somewhat complicated, in the “canonical” special case where the sender’s utility is independent of the state we obtain the simple result that single-dipped disclosure is optimal when the cross-partial derivative of the receiver’s utility (with respect to action and state) is log-supermodular, while single-peaked disclosure is optimal when it is log-submodular. Intuitively, maximizing the receiver’s action corresponds to maximizing his expected marginal utility (from increasing his action). Thus, whether more dispersed states should be allocated to higher or lower actions depends on whether the receiver’s marginal utility is more convex in the state at higher or lower actions. More convex marginal utility at higher actions corresponds to a log-supermodular

cross-partial, in which case single-dipped disclosure is optimal (i.e., dispersed states are allocated to higher actions). More convex marginal utility at lower actions corresponds to a log-submodular cross-partial, in which case single-peaked disclosure is optimal (i.e., dispersed states are allocated to lower actions). We prove these results via simple perturbation arguments that consider the consequences of re-allocating probability mass over only three states to each of two actions, where mass on the center state is re-allocated from one action to the other, while mass on the extreme states is re-allocated in the opposite direction.

The canonical case covers the [Friedman and Holden](#) and fund manager applications. Even in non-canonical environments, where our general characterization is more complicated, our analysis of applications (like the [Goldstein and Leitner](#) and [Guo and Shmaya](#) settings) seems much simpler than previous treatments, and is again based on simple re-allocations of probability mass over three states to each of two actions.

2. MODEL

We consider a standard persuasion problem with one sender and one receiver. The sender's utility $V(\rho, \theta)$ and the receiver's utility $U(\rho, \theta)$ depend on the receiver's action $\rho \in [0, 1]$ and the state of the world $\theta \in [0, 1]$. The sender and receiver share a common prior ϕ over the state space.

The sender chooses a signal that reveals information about the state and thus induces a distribution τ of the posterior belief μ . By a well-known ‘‘splitting lemma’’ of [Aumann and Maschler \(1995\)](#) and [Kamenica and Gentzkow \(2011\)](#), there exists a signal that induces distribution τ if and only if the average posterior equals the prior, so that $\int \mu d\tau = \phi$. We simply refer to such a distribution τ as a signal.

The timing is as follows: The sender chooses a signal τ . The posterior belief μ is drawn according to τ . The receiver takes an action ρ that maximizes his expected utility $\int U(\rho, \theta) d\mu$. Whenever the receiver has multiple optimal actions, he breaks ties in favor of the sender. An *optimal* signal is one that maximizes the sender's expected utility.²

We impose three standard assumptions on the utility functions. First, we assume utilities are smooth.

Assumption 1. $V(\rho, \theta)$ and $U(\rho, \theta)$ are absolutely continuous in ρ , and the marginal utilities

$$v(\rho, \theta) = \frac{\partial V(\rho, \theta)}{\partial \rho} \quad \text{and} \quad u(\rho, \theta) = \frac{\partial U(\rho, \theta)}{\partial \rho}$$

are continuous in θ .

Second, we assume the sender strictly prefers higher actions.

²Formally, the sender's problem is to find τ that satisfies $\int \mu d\tau = \phi$ and maximizes $\int \widehat{V}(\mu) d\tau$, where $\widehat{V}(\mu) = \max_{\rho \in \rho^*(\mu)} \int V(\rho, \theta) d\mu$ and $\rho^*(\mu) = \arg \max_{\rho \in [0, 1]} \int U(\rho, \theta) d\mu$.

Assumption 2. $V(\rho, \theta)$ is strictly increasing in ρ .

Third, we assume the receiver prefers higher actions at higher states, and his expected utility is single-peaked in his action for any posterior belief.

Assumption 3. $u(\rho, \theta)$ satisfies *upcrossing* in θ : for all $\theta < \theta'$,

$$u(\rho, \theta) \geq (>)0 \implies u(\rho, \theta') \geq (>)0.$$

Furthermore, $u(\rho, \theta)$ satisfies *aggregate downcrossing* in ρ : for all μ and all $\rho < \rho'$,

$$\int u(\rho', \theta) d\mu \geq (>)0 \implies \int u(\rho, \theta) d\mu \geq (>)0.$$

Aggregate downcrossing has been characterized by [Quah and Strulovici \(2012\)](#) and [Choi and Smith \(2017\)](#). For completeness, we include a simple, self-contained proof of their results in [Appendix A.1](#).

Under our tie-breaking assumption, the posterior belief μ induces action

$$\rho^*(\mu) = \max \left\{ \rho : \int u(\rho, \theta) d\mu \geq 0 \right\}. \quad (1)$$

Indeed, the receiver's expected utility is (weakly) increasing in ρ for $\rho \leq \rho^*(\mu)$ and strictly decreasing in ρ for $\rho > \rho^*(\mu)$. Thus, $\rho^*(\mu)$ is the receiver's highest optimal action. Since the sender strictly prefers higher actions, the receiver breaks ties in her favor by taking $\rho^*(\mu)$. Intuitively, (1) implies that, to maximize the receiver's action, the sender must maximize his expected marginal utility.

An important special case of the model arises when the sender's utility is state-independent, so that V depends only on ρ , and the receiver's utility is strictly super-modular, so that $\partial u / \partial \theta > 0$.³ We call this the *canonical case*. Many of our results will be particularly easy to interpret in this case. When we turn to applications in [Section 4](#), those considered in [Sections 4.1](#) and [4.2](#) are canonical, while the ones in [Sections 4.3](#) through [4.5](#) are not.

Our model implicitly allows the sets of states and actions to be finite. The finite-states case arises when the support of ϕ , denoted by $\text{supp}(\phi)$, is a finite set. The finite-actions case arises when the utility functions U and V are piecewise linear in ρ , with each piece representing expected utilities spanned by lotteries on two consecutive actions.

3. GENERAL RESULTS

3.1. Pairwise Signals. We first show that, without loss of generality, we can restrict attention to signals that induce posterior distributions with at most binary support. A signal τ is *pairwise* if each $\mu \in \text{supp}(\tau)$ takes the form $\mu = p(\mu)\delta_{\theta(\mu)} + (1-p(\mu))\delta_{\theta'(\mu)}$ for some states $\theta(\mu) \leq \theta'(\mu)$ and some probability $p(\mu) \in [0, 1]$, where δ_θ denotes the degenerate distribution that assigns probability one to state θ .

³Here u is implicitly required to be differentiable in θ , so that $\partial u / \partial \theta$ is well-defined.

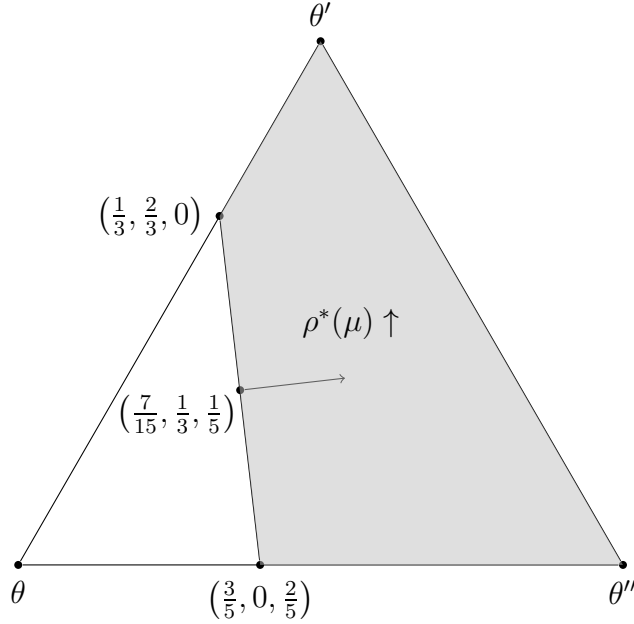


FIGURE 1. Pairwise Signal

Notes: The arrow shows the direction of fastest increase of the receiver's marginal utility. The shaded area shows posteriors that induce weakly higher actions than $\rho^*(\mu)$. The extreme points of the “indifference line” in the simplex have binary support.

Lemma 1. *There exists an optimal signal that is pairwise.*

The intuition is that, for any posterior distribution, there exists a hyperplane passing through it such that all distributions on the hyperplane induce a weakly higher receiver action, and the extreme points of the hyperplane in the simplex have at most binary support. Thus, any posterior that puts weight on more than two states can be split into posteriors with at most binary support, each of which induces a weakly higher action, which is preferred by the sender.

For a posterior that puts weight on three states, the argument can be seen graphically.

Example 1. Suppose that $\text{supp}(\mu) = \{\theta, \theta', \theta''\}$, $(\mu(\theta), \mu(\theta'), \mu(\theta'')) = (7/15, 1/3, 1/5)$, $(u(1/2, \theta), u(1/2, \theta'), u(1/2, \theta'')) = (-2, 1, 3)$, and $\rho^*(\mu) = 1/2$, so that

$$\int u(\rho^*(\mu), \theta) d\mu = -2 \cdot \frac{7}{15} + 1 \cdot \frac{1}{3} + 3 \cdot \frac{1}{5} = 0.$$

Consider a signal that splits μ into two equiprobable posteriors $(\mu_1(\theta), \mu_1(\theta'), \mu_1(\theta'')) = (1/3, 2/3, 0)$ and $(\mu_2(\theta), \mu_2(\theta'), \mu_2(\theta'')) = (3/5, 0, 2/5)$. See Figure 1. This signal is feasible because

$$\frac{1}{2}\mu_1 + \frac{1}{2}\mu_2 = \frac{1}{2} \cdot \left(\frac{1}{3}, \frac{2}{3}, 0\right) + \frac{1}{2} \cdot \left(\frac{3}{5}, 0, \frac{2}{5}\right) = \left(\frac{7}{15}, \frac{1}{3}, \frac{1}{5}\right) = \mu.$$

Moreover, $\rho^*(\mu_1) \geq \rho^*(\mu)$ and $\rho^*(\mu_2) \geq \rho^*(\mu)$ because

$$\int u(\rho^*(\mu), \theta) d\mu_1 = -2 \cdot \frac{1}{3} + 1 \cdot \frac{2}{3} = 0,$$

$$\int u(\rho^*(\mu), \theta) d\mu_2 = -2 \cdot \frac{3}{5} + 3 \cdot \frac{2}{5} = 0.$$

Perhaps surprisingly, we show that pooling more than two states is generally sub-optimal. Specifically, Theorem 1 shows that an optimal signal never generates a posterior belief that (i) induces an interior action, (ii) contains three distinct states in its support, and (iii) satisfies a non-singularity condition involving the utility functions at the induced action and the three states. For simplicity, we present the result for the case where ϕ has finite support.

Theorem 1. *Let ϕ have finite support. There does not exist an optimal signal τ and a belief μ with $\tau(\mu) > 0$ such that (i) $\rho^*(\mu) \in (0, 1)$, (ii) $(\mu(\theta), \mu(\theta'), \mu(\theta'')) > (0, 0, 0)$ for $\theta < \theta' < \theta''$, and (iii) V and u are continuously differentiable in ρ in a neighborhood of $\rho^*(\mu)$, $\int (\partial u(\rho^*(\mu), \theta) / \partial \rho) d\mu \neq 0$, and*

$$\begin{vmatrix} v(\rho^*(\mu), \theta) & v(\rho^*(\mu), \theta') & v(\rho^*(\mu), \theta'') \\ u(\rho^*(\mu), \theta) & u(\rho^*(\mu), \theta') & u(\rho^*(\mu), \theta'') \\ \frac{\partial u(\rho^*(\mu), \theta)}{\partial \rho} & \frac{\partial u(\rho^*(\mu), \theta')}{\partial \rho} & \frac{\partial u(\rho^*(\mu), \theta'')}{\partial \rho} \end{vmatrix} \neq 0. \quad (2)$$

The conclusion of Theorem 1 differs starkly from that of a large number of papers that focus on the case where utilities are linear in the state and emphasize the optimality of pooling intervals of states: see Section 4.1 for references. Theorem 1 reveals that the optimality of interval pooling in these papers rests on the “non-genericity” of linear utility.

To see the intuition for Theorem 1, consider a posterior distribution μ with $\text{supp}(\mu) = \{\theta, \theta', \theta''\}$ and interior $\rho^*(\mu)$, where $\theta < \theta' < \theta''$. Lemma 1 implies that we can split μ into posterior distributions μ_1 and μ_2 such that μ_1 and μ_2 induce action $\rho^*(\mu)$ or higher and both posteriors have at most two states in the support. For example, suppose that $\text{supp}(\mu_1) = \{\theta, \theta'\}$ and $\text{supp}(\mu_2) = \{\theta, \theta''\}$. Consider a perturbation that moves a probability mass dp at θ from μ_1 to μ_2 .⁴ This perturbation induces non-zero marginal changes in the receiver’s action at μ_1 and μ_2 . In the non-singular case, these changes have a non-zero marginal effect on the sender’s expected utility by the Implicit Function Theorem. Therefore, either this perturbation or the reverse perturbation, where dp is replaced with $-dp$, is strictly profitable for the sender.

⁴More formally, moving a probability mass dp at θ from μ_1 to μ_2 corresponds to changing a signal τ whose support contains μ_1 and μ_2 in the following way. For simplicity, suppose that $\text{supp}(\phi)$ is a finite set and $\text{supp}(\tau) = \{\mu_1, \mu_2\}$. First, change $\tau(\mu_1)$ to $\tau'(\mu'_1) = \tau(\mu_1) - dp$ where μ'_1 satisfies $\mu'_1(\theta)\tau'(\mu'_1) = \mu_1(\theta)\tau(\mu_1) - dp$ and $\mu'_1(\theta')\tau'(\mu'_1) = \mu_1(\theta')\tau(\mu_1)$ for all $\theta' \neq \theta$. Second, change $\tau(\mu_2)$ to $\tau'(\mu'_2)$ in a similar way.

This argument requires the induced action $\rho^*(\mu)$ to be interior. Induced actions are always interior if the marginal utilities v and u satisfy Inada-type conditions.⁵

In light of Lemma 1 and Theorem 1, we henceforth restrict attention to pairwise signals. We can thus represent a posterior μ as a pair of states $\theta, \theta' \in \text{supp}(\phi)$ and a probability p of θ . Note that posterior μ is non-degenerate if $\theta \neq \theta'$ and $p \in (0, 1)$, and is degenerate otherwise. Unless otherwise specified, given posteriors $\mu_1, \mu_2 \in \text{supp}(\tau)$, we denote (and order) their supports by $\theta_1 \leq \theta'_1$ and $\theta_2 \leq \theta'_2$, respectively.

Given this restriction to pairwise signals, the sender's problem is to choose what pairs of states $\{\theta, \theta'\}$ to pool and what probabilities $\{p, 1 - p\}$ to assign to states $\{\theta, \theta'\}$ to generate posterior beliefs $p\delta_\theta + (1 - p)\delta_{\theta'}$. This problem generalizes standard matching problems where, as in Becker (1973), each state in a pair $\{\theta, \theta'\}$ is required to have equal weight, so that $p = 1/2$. Following Becker, Legros and Newman (2002) characterize conditions under which simple matching patterns are optimal, focusing on variants of “positively assortative” and “negatively assortative” matching.

We aim to characterize conditions under which simple disclosure patterns are optimal. We focus on *assortative disclosure*, which requires the supports of any two posterior distributions not to cross, so that all pairs of supports are either ordered or nested. Specifically, a pairwise signal τ is *assortative* if for all $\mu_1, \mu_2 \in \text{supp}(\tau)$, we have

$$\text{neither } \theta_1 < \theta_2 < \theta'_1 < \theta'_2 \text{ nor } \theta_2 < \theta_1 < \theta'_2 < \theta'_1.$$

If in addition all pairs of supports are nested, so that for all $\mu_1, \mu_2 \in \text{supp}(\tau)$, we have

$$\text{either } \theta_1 \leq \theta_2 \leq \theta'_2 \leq \theta'_1 \text{ or } \theta_2 \leq \theta_1 \leq \theta'_1 \leq \theta'_2,$$

we then say that assortative disclosure is *unitary*. A unitary signal thus matches all states within a single “block.” See Figure 2.

We characterize three special cases of assortative disclosure: (i) full disclosure, where the sender fully separates states; (ii) single-dipped disclosure, where ρ is single-dipped on sets of nested pairs of states; and (iii) single-peaked disclosure, where ρ is single-peaked on sets of nested pairs of states.

3.2. Full Disclosure. *Full disclosure* is a signal τ that fully separates each state θ , so that $\mu \in \text{supp}(\tau)$ if and only if $\mu = \delta_\theta$ for some θ .⁶

⁵Otherwise, an optimal signal can induce action 0 at many states. All such states θ , however, must satisfy $u(0, \theta) \leq 0$, so that the sender does not benefit from separating these states. An optimal signal can also induce action 1 at many states. In particular, if $\int u(1, \theta)d\phi \geq 0$, then a completely uninformative signal induces action 1 and thus is optimal; otherwise, if $\rho^*(\mu) = 1$ then $\int u(1, \theta)d\mu = 0$, so the sender does not benefit from pooling more states to induce action 1 with higher probability.

⁶In the case where the prior distribution ϕ admits a density, full disclosure can be equivalently defined as *positive assortative disclosure* that requires the supports of posterior distributions to be ordered, so that for all $\mu_1, \mu_2 \in \text{supp}(\tau)$, we have either $\theta_1 \leq \theta'_1 \leq \theta_2 \leq \theta'_2$ or $\theta_2 \leq \theta'_2 \leq \theta_1 \leq \theta'_1$. Somewhat counterintuitively, full disclosure can also be represented as unitary assortative disclosure, where each posterior $\mu \in \text{supp}(\tau)$ takes the form $\mu = 0\delta_0 + 1\delta_\theta$ for some $\theta \in [0, 1]$.

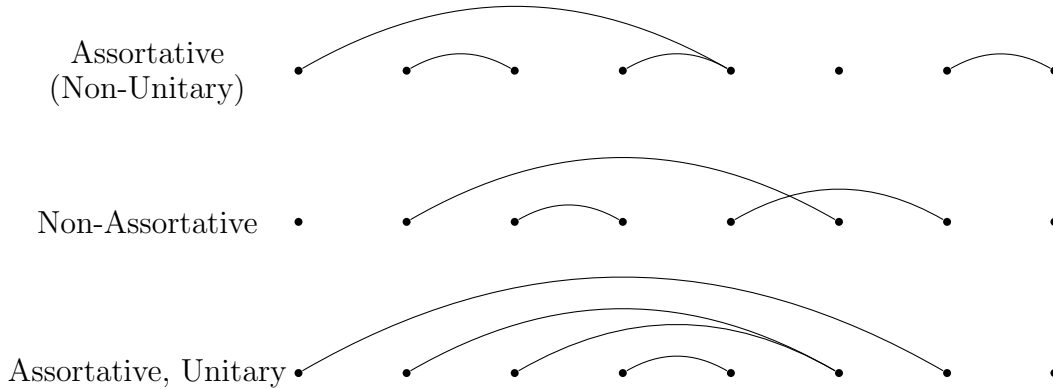


FIGURE 2. Assortative, Non-Assortative, and Unitary Signals

Notes: Links show supports of posteriors induced by a pairwise signal. The top signal is assortative because paired states do not cross but is not unitary because the first and last pairs are ordered. The middle signal is not assortative because the first and last pairs cross. The bottom signal is assortative and unitary, because all paired states are nested (although the last state is separated, it can be viewed as being pooled with the first state with zero probability).

If, for all states θ and θ' , and all probabilities p , the sender prefers to split the posterior $\mu = p\delta_\theta + (1-p)\delta_{\theta'}$ into degenerate posteriors δ_θ and $\delta_{\theta'}$, then the sender prefers full disclosure to any pairwise signal. Since restricting to pairwise signals is without loss of generality, we conclude that full disclosure is optimal in this case.

Conversely, if the sender strictly prefers not to split $\mu = p\delta_\theta + (1-p)\delta_{\theta'}$ into δ_θ and $\delta_{\theta'}$ for some states $\theta, \theta' \in \text{supp}(\phi)$ and some probability p , then the sender strictly prefers a pairwise signal that differs from full disclosure only in that it pools states θ and θ' into μ ; so full disclosure is not optimal.⁷ Recalling that belief $\mu = p\delta_\theta + (1-p)\delta_{\theta'}$ induces action

$$\rho^*(\mu) = \max\{\rho : pu(\rho, \theta) + (1-p)u(\rho, \theta') \geq 0\},$$

we obtain the following result.⁸

Theorem 2. *Full disclosure is optimal if and only if for all $\mu = p\delta_\theta + (1-p)\delta_{\theta'}$, with $\theta, \theta' \in \text{supp}(\phi)$ and $p \in [0, 1]$,*

$$pV(\rho^*(\delta_\theta), \theta) + (1-p)V(\rho^*(\delta_{\theta'}), \theta') \geq pV(\rho^*(\mu), \theta) + (1-p)V(\rho^*(\mu), \theta'). \quad (3)$$

In the canonical case, where V does not depend on θ , condition (3) simplifies as follows:

⁷This argument is valid when ϕ has finite support; the general case is covered in the Appendix.

⁸Kolotilin (2018, Proposition 1, part 2) obtains a similar result using duality theory, under the additional assumption that the prior admits a strictly positive density.

Corollary 2. *In the canonical case, full disclosure is optimal if and only if for all $\mu = p\delta_\theta + (1-p)\delta_{\theta'}$, with $\theta, \theta' \in \text{supp}(\phi)$ and $p \in [0, 1]$,*

$$pV(\rho^*(\delta_\theta)) + (1-p)V(\rho^*(\delta_{\theta'})) \geq V(\rho^*(\mu)). \quad (4)$$

Becker (1973) famously showed that, if the utility from matching two types $h(\theta, \theta')$ is supermodular, then it is optimal to match like types. Legros and Newman (2002) refer to this extreme form of positive assortative matching as *segregation*. Their Propositions 4 and 9 show that segregation is optimal if and only if $h(\theta, \theta) + h(\theta', \theta') \geq 2h(\theta, \theta')$ for all θ, θ' (which is a strictly weaker property than supermodularity). In the context of persuasion, segregation corresponds to full disclosure. Indeed, if we fix $p = 1/2$ and let $h(\theta, \theta') = V(\rho^*(\delta_\theta/2 + \delta_{\theta'}/2))$, then (4) reduces to precisely Legros and Newman's condition.

Another interesting special case arises when the receiver's utility is linear in θ . Following Gentzkow and Kamenica (2016), in this case it is without loss of generality to normalize $\rho^*(\mu) = \mathbb{E}_\mu[\theta]$. Theorem 2 then yields the following result.

Corollary 2'. *Let $\rho^*(\mu) = \mathbb{E}_\mu[\theta]$. Full disclosure is optimal if $V(\rho, \theta)$ is convex in ρ and satisfies $V(\theta, \theta) + V(\theta', \theta') \geq V(\theta, \theta') + V(\theta', \theta)$ for all $\theta, \theta' \in \text{supp}(\phi)$.*

Proof. Condition (3) holds because

$$\begin{aligned} & pV(p\theta + (1-p)\theta', \theta) + (1-p)V(p\theta + (1-p)\theta', \theta') \\ & \leq p(pV(\theta, \theta) + (1-p)V(\theta', \theta)) + (1-p)(pV(\theta, \theta') + (1-p)V(\theta', \theta')) \\ & \leq pV(\theta, \theta) + (1-p)V(\theta', \theta'), \end{aligned}$$

where the first inequality holds because $V(\rho, \theta)$ is convex in ρ , and the second holds because $V(\theta, \theta) + V(\theta', \theta') \geq V(\theta, \theta') + V(\theta', \theta)$. \square

A sufficient condition for $V(\theta, \theta) + V(\theta', \theta') \geq V(\theta, \theta') + V(\theta', \theta)$ is supermodularity of V : for all $\theta < \theta'$ and $\rho < \rho'$, $V(\rho, \theta) + V(\rho', \theta') \geq V(\rho, \theta') + V(\rho', \theta)$. Thus, full disclosure is optimal whenever the receiver's utility is linear in θ and the sender's utility is convex in ρ and supermodular in (ρ, θ) . This sufficient condition for full disclosure generalizes that given by Rayo and Segal (2010, Lemma 11).

3.3. Single-Dipped Disclosure. *Single-dipped disclosure* is a pairwise signal τ such that for each $\mu_1, \mu_2 \in \text{supp}(\tau)$ that induce actions $\rho_1 < \rho_2$, we have

$$\theta_2, \theta'_2 \notin (\theta_1, \theta'_1).$$

That is, single-dipped disclosure is assortative disclosure where, within each set of nested pairs of states, the receiver's action is higher at more extreme states. If single-dipped disclosure is unitary, then the receiver's action is single-dipped on $[0, 1]$.

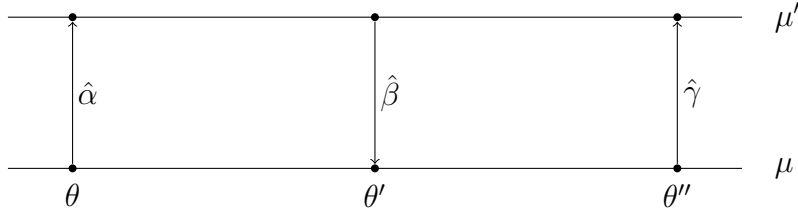


FIGURE 3. A Profitable Perturbation for a Non-Single-Dipped Signal

Notes: The figure shows a perturbation of a signal that shifts weights $\hat{\alpha}$ and $\hat{\gamma}$ on θ and θ'' from μ to μ' and shifts weight $\hat{\beta}$ on θ' from μ' to μ . This perturbation is profitable if it increases the induced actions ρ and ρ' and also increases the sender's expected utility for fixed ρ and ρ' .

To simplify notation, we define, for $\theta < \theta' < \theta''$ and $\rho < \rho'$,

$$a = \frac{u(\rho', \theta)}{V(\rho', \theta) - V(\rho, \theta)}, \quad b = \frac{u(\rho', \theta')}{V(\rho', \theta') - V(\rho, \theta')}, \quad c = \frac{u(\rho', \theta'')}{V(\rho', \theta'') - V(\rho, \theta'')},$$

$$A = \frac{u(\rho, \theta)}{V(\rho', \theta) - V(\rho, \theta)}, \quad B = \frac{u(\rho, \theta')}{V(\rho', \theta') - V(\rho, \theta')}, \quad C = \frac{u(\rho, \theta'')}{V(\rho', \theta'') - V(\rho, \theta'')}.$$

Since V is strictly increasing in ρ , the denominators in the above six equations are strictly positive. Hence, since u satisfies upcrossing in θ , we have

$$a \geq 0 \implies b \geq 0, \quad b \geq 0 \implies c \geq 0,$$

$$A \geq 0 \implies B \geq 0, \quad B \geq 0 \implies C \geq 0.$$

Similarly, since u satisfies downcrossing in ρ , we have

$$a \geq 0 \implies A \geq 0, \quad b \geq 0 \implies B \geq 0, \quad c \geq 0 \implies C \geq 0.$$

Lemma 3 is our key tool for verifying optimality of single-dipped disclosure. We will use this lemma to establish necessary and sufficient conditions on the utility functions for single-dipped disclosure to be optimal (Theorem 3), but in the applications in Section 4 we find it is often easier to apply the lemma directly.

Lemma 3. *Single-dipped disclosure is optimal if, for all $\theta < \theta' < \theta''$ and $\rho < \rho'$ such that $A < 0 < C$, there exists a non-zero vector $(\alpha, \beta, \gamma) \geq 0$ such that*

$$\begin{pmatrix} a & -b & c \\ -A & B & -C \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} \geq \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \quad (5)$$

Figure 3 illustrates the proof of Lemma 3. Let μ and μ' denote posteriors in the support of a signal τ inducing actions ρ and ρ' , respectively. When the conditions of the lemma are satisfied, a perturbation of signal τ that shifts (probability) weight $\hat{\alpha}$ on θ from μ to μ' , shifts weight $\hat{\beta}$ on θ' from μ' to μ , and shifts weight $\hat{\gamma}$ on θ''

from μ to μ' increases both induced actions ρ and ρ' , and also increases the sender's expected utility for fixed ρ and ρ' , where

$$\hat{\alpha} = \frac{\alpha}{V(\rho', \theta) - V(\rho, \theta)}, \quad \hat{\beta} = \frac{\beta}{V(\rho', \theta') - V(\rho, \theta')}, \quad \hat{\gamma} = \frac{\gamma}{V(\rho', \theta'') - V(\rho, \theta'')}.$$

Repeatedly applying such a perturbation yields single-dipped disclosure.⁹ One may also establish a partial converse to Lemma 3.¹⁰

Lemma 3 easily yields sufficient and “almost necessary” conditions on the utility functions for single-dipped disclosure to be optimal.

Theorem 3. *Single-dipped disclosure is optimal if, for all $\theta < \theta' < \theta''$ and $\rho < \rho'$ such that $A < 0 < C$, we have*

$$\begin{aligned} Ab &\geq Ba, & Bc &\geq Cb, & Ac &\geq Ca, \\ Ab - Ba + Bc - Cb &\geq Ac - Ca. \end{aligned} \tag{6}$$

Conversely, (6) holds for $\theta < \theta' < \theta''$ and $\rho < \rho'$ such that $A < 0 < C$ and $c > 0$ only if there exists a non-zero vector $(\alpha, \beta, \gamma) \geq 0$ that satisfies (5).

Proof. Consider a vector

$$(\alpha, \beta, \gamma) = \begin{cases} (Bc - Cb, Ac - Ca, Ab - Ba), & \text{if } Ac > Ca, \\ (C, 0, -A), & \text{if } Ac = Ca. \end{cases}$$

Consider first the case $Ac > Ca$. The first line in (6) implies that $(\alpha, \beta, \gamma) \geq 0$. That is, a perturbation in Figure 3 is well-defined. The first and second inequalities in (5) hold with equality. That is, the perturbation holds ρ and ρ' fixed. Moreover, the third inequality in (5) holds because it is equivalent to the second line in (6). That is, the perturbation increases the sender's expected utility for fixed ρ and ρ' .

Consider now the case $Ac = Ca$. The inequality $A < 0 < C$ implies that $(\alpha, \beta, \gamma) \geq 0$. The first and second inequalities in (5) hold with equality. Moreover, the third inequality in (5) holds because $A < 0 < C$. Again, the perturbation holds ρ and ρ' fixed and increases the sender's expected utility.

Since (5) holds in both cases, single-dipped disclosure is optimal by Lemma 3.

Finally, the converse says that the perturbation that holds ρ and ρ' fixed and increases the sender's expected utility yields not only sufficient but also necessary conditions

⁹If it is possible to satisfy (5) with at least one strict inequality, this argument immediately implies that all optimal signals are single-dipped disclosure. Otherwise, the perturbation gives only a weak improvement, and our proof is fully rigorous only when ϕ has finite support.

¹⁰If the sets of states and actions are finite and the utility functions V and U do not satisfy (5) for some $\theta < \theta' < \theta''$ and $0 < \rho < \rho'$ such that $A < 0 < C$, then there exist a prior ϕ' and utility functions V' and U' that coincide with V and U for $(\rho, \theta) \in \{\rho, \rho'\} \times \{\theta, \theta', \theta''\}$ and satisfy (5) for all other (ρ, θ) such that single-dipped disclosure is suboptimal.

for single-dipped disclosure to be optimal in the case where actions ρ and ρ' are sufficiently close to each other, so that

$$u(\rho, \theta'') > 0 \implies u(\rho', \theta'') > 0.$$

We prove this converse in the Appendix. It is possible to provide necessary and sufficient conditions for (5) to hold not only in the case $c > 0$ but also in the case $c \leq 0$, but since these conditions are quite complicated and are not used in our applications, we omit them. \square

In the canonical case, condition (6) simplifies as follows:

Corollary 3. *In the canonical case, single-dipped disclosure is optimal if $\partial u / \partial \theta$ is log-supermodular: for all $\theta < \theta'$ and $\rho < \rho'$,*

$$\frac{\partial u(\rho, \theta)}{\partial \theta} \frac{\partial u(\rho', \theta')}{\partial \theta} \geq \frac{\partial u(\rho, \theta')}{\partial \theta} \frac{\partial u(\rho', \theta)}{\partial \theta}. \quad (7)$$

For a simple geometric intuition, fix $\theta < \theta' < \theta''$, and suppose there exist posteriors $\mu, \mu' \in \text{supp}(\tau)$ inducing actions ρ and ρ' , with $\theta, \theta'' \in \text{supp}(\mu)$, $\theta' \in \text{supp}(\mu')$, and $\rho < \rho'$. With state-independent sender utility, this signal cannot be optimal if we can shift weight on θ and θ'' from μ to μ' and shift an equal (or lesser) weight on θ' from μ' to μ such that both ρ and ρ' increase. When u is strictly increasing in θ , such a perturbation exists if and only if the ratio of the slopes of the line segments formed by linearly interpolating between $u(\rho, \theta)$ and $u(\rho, \theta')$, and between $u(\rho, \theta')$ and $u(\rho, \theta'')$, is greater than the corresponding ratio with ρ' in place of ρ : indeed, this is precisely the condition that, when maximizing the receiver's expected marginal utility, the marginal rate of substitution (MRS) between “shifting weight from θ to θ'' ” and “shifting weight from θ' to θ'' ” is higher with action ρ than with action ρ' . See Figure 4. Letting \tilde{u} be the interpolated value of u , this condition is

$$\frac{\partial \tilde{u}(\rho, \tilde{\theta})}{\partial \theta} \Big/ \frac{\partial \tilde{u}(\rho, \tilde{\theta}')}{\partial \theta} \geq \frac{\partial \tilde{u}(\rho', \tilde{\theta})}{\partial \theta} \Big/ \frac{\partial \tilde{u}(\rho', \tilde{\theta}')}{\partial \theta} \quad \text{for all } \tilde{\theta} \in (\theta, \theta'), \tilde{\theta}' \in (\theta', \theta'').$$

When u is differentiable in θ , this condition holds for any choice of $\theta < \theta' < \theta''$ if

$$\frac{\partial u(\rho, \theta)}{\partial \theta} \Big/ \frac{\partial u(\rho, \theta')}{\partial \theta} \geq \frac{\partial u(\rho', \theta)}{\partial \theta} \Big/ \frac{\partial u(\rho', \theta')}{\partial \theta} \quad \text{for all } \theta < \theta',$$

which is equivalent to (7), since u is strictly increasing in θ in the canonical case.

Note also that another way of writing the MRS condition is

$$\frac{u(\rho, \theta') - u(\rho, \theta)}{u(\rho, \theta'') - u(\rho, \theta')} \geq \frac{u(\rho', \theta') - u(\rho', \theta)}{u(\rho', \theta'') - u(\rho', \theta')},$$

or

$$\frac{B - A}{C - B} \geq \frac{b - a}{c - b}.$$

Since u is strictly increasing in θ in the canonical case, so that $C > B$ and $c > b$, this is equivalent to

$$Ab - Ba + Bc - Cb \geq Ac - Ca.$$

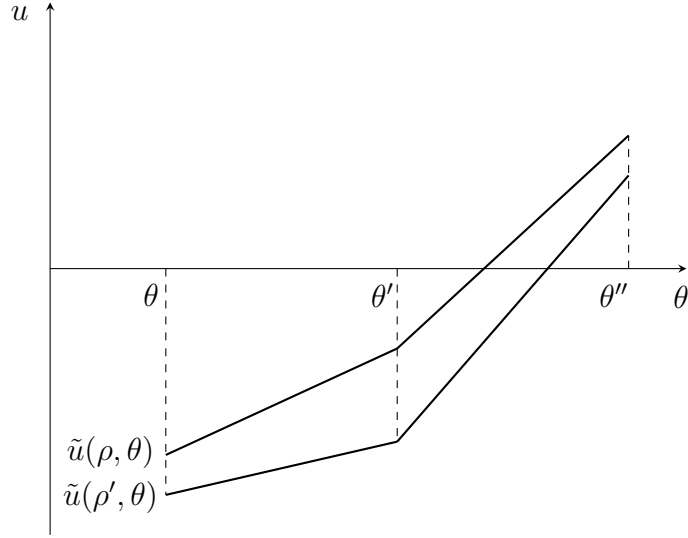


FIGURE 4. Log-Supermodular $\partial u/\partial\theta$ and Single-Dipped τ

Notes: The receiver’s marginal utility is more convex in the state at higher actions. More extreme states should thus induce higher actions.

Theorem 3 shows that the same condition plays a key role in characterizing optimality of single-dipped disclosure beyond the canonical case.

A close antecedent to Corollary 3, albeit in a specific and somewhat unexpected context, is Friedman and Holden (2008, Lemma 1). This result shows that, under a condition called the “informative signal property,” if a politician designing an optimally gerrymandered districting plan allocates two voter types to the same district, then all voter types in between these two must be allocated to districts in which the median voter is less favorable. We investigate the connection between optimal gerrymandering and assortative information disclosure in a companion paper, where we confirm that this property is equivalent to single-dippedness, and that the informative signal property is equivalent to log-supermodularity of u .

We now sketch how an optimal single-dipped disclosure signal can be found using an optimal control approach. We consider the unitary case, which is relevant in many special cases, as shown in Section 4.

Suppose that ϕ has a strictly positive density f . Consider a single-dipped disclosure signal τ . This signal can be represented by a cutoff state $\theta^* \in [0, 1]$ and an absolutely continuous decreasing matching function $n : [\theta^*, 1] \rightarrow [0, \theta^*]$ with $n(1) = 0$ and $n(\theta^*) = \theta^*$. Each $\mu \in \text{supp}(\tau)$ assigns probabilities $p(\theta)$ and $1 - p(\theta)$ to θ and $n(\theta)$ with $\theta \in [\theta^*, 1]$ and

$$p(\theta) = \frac{f(\theta)}{f(\theta) + f(n(\theta)) \left| \frac{dn(\theta)}{d\theta} \right|}.$$

In particular, state $\theta \in [\theta^*, 1]$ is separated if and only if $dn(\theta)/d\theta = 0$, but state $\theta \in [0, \theta^*)$ cannot be separated, otherwise the induced action would not be single-dipped on $[0, 1]$.¹¹ The induced action for states $\theta \in [\theta^*, 1]$ and $n(\theta) \in [0, \theta^*]$ is

$$\rho^*(\theta) = \rho^*(n(\theta)) = \max\{\rho : u(\rho, \theta)p(\theta) + u(\rho, n(\theta))(1 - p(\theta)) \geq 0\},$$

which is decreasing in θ on $[0, \theta^*]$ and increasing in θ on $[\theta^*, 1]$ for unitary single-dipped disclosure. The optimal cutoff and matching function maximize the sender's expected utility

$$\int_{\theta^*}^1 \left[V(\rho^*(\theta), \theta)f(\theta) + V(\rho^*(\theta), n(\theta))f(n(\theta)) \left| \frac{dn(\theta)}{d\theta} \right| \right] d\theta.$$

3.4. Single-Peaked Disclosure. *Single-peaked disclosure* is a pairwise signal τ such that for each $\mu_1, \mu_2 \in \text{supp}(\tau)$ that induce actions $\rho_1 < \rho_2$, we have

$$\theta_1, \theta'_1 \notin (\theta_2, \theta'_2).$$

That is, single-peaked disclosure is assortative disclosure where, within each set of nested pairs of states, the receiver's action is lower at more extreme states. If single-peaked disclosure is unitary, then the receiver's action is single-peaked on $[0, 1]$.

Our conditions for the optimality of single-peaked disclosure are almost completely symmetric to those for single-dipped disclosure, where now the MRS condition takes the opposite sign.

Lemma 4. *Single-peaked disclosure is optimal if, for all $\theta < \theta' < \theta''$ and $\rho < \rho'$ such that $a < 0 < c$, there exists a non-zero vector $(\alpha, \beta, \gamma) \geq 0$ such that*

$$\begin{pmatrix} a & -b & c \\ -A & B & -C \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} \leq \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \quad (8)$$

We omit the proof of Lemma 4, which is parallel to that of Lemma 3. The relevant perturbation simply reverses the direction of the arrows in Figure 3: if $\mu, \mu' \in \text{supp}(\tau)$ induce actions $\rho < \rho'$ and $\theta' \in \text{supp}(\mu)$ lies between $\theta, \theta'' \in \text{supp}(\mu')$ (so disclosure is not single-peaked), then τ can be improved by shifting weight $\hat{\beta}$ on θ' from μ to μ' , while shifting weight $\hat{\alpha}$ and $\hat{\gamma}$ on θ and θ'' from μ' to μ .

Lemma 4 easily yields conditions on the utility functions for single-dipped disclosure to be optimal.

Theorem 4. *Single-peaked disclosure is optimal if, for all $\theta < \theta' < \theta''$ and $\rho < \rho'$ such that $a < 0 < c$, we have*

$$\begin{aligned} a \leq b < c \text{ or } A < B \leq C, \\ Ab - Ba + Bc - Cb \leq Ac - Ca. \end{aligned} \quad (9)$$

¹¹The extreme case $\theta^* = 0$ is just full disclosure in disguise, where each posterior $\mu \in \text{supp}(\tau)$ takes the form $\mu = 0\delta_0 + 1\delta_\theta$ for some $\theta \in [0, 1]$.

Conversely, (9) holds for $\theta < \theta' < \theta''$ and $\rho < \rho'$ such that $a < 0 < c$ and $A < 0$ only if there exists a non-zero vector $(\alpha, \beta, \gamma) \geq 0$ that satisfies (8).

The second line of (9) is the same MRS condition as the second line of (6), with the opposite sign. If the first line of (9) were replaced with the first line of (6), the sufficiency part of Theorem 4 would generically hold true,¹² but the necessity part would fail. This reflects a subtle difference in the “binding” perturbation that we must consider to obtain necessary and sufficient conditions for single-dipped and single-peaked disclosure. In the single-dipped case, the binding perturbation holds ρ and ρ' fixed while increasing the sender’s expected utility. As the proof of Theorem 3 shows, such a perturbation exists whenever the MRS condition and the first line of (6) hold. In the single-peaked case, the binding perturbation holds one of the two actions ρ and ρ' fixed and also fixes the sender’s expected utility (for fixed ρ and ρ'), while increasing the other action. The proof of Theorem 4 shows that such a perturbation exists whenever the (reversed) MRS condition and the first line of (9) hold.

In the canonical case, condition (9) simplifies as follows:

Corollary 4. *In the canonical case, single-peaked disclosure is optimal if $\partial u / \partial \theta$ is log-submodular: for all $\theta < \theta'$ and $\rho < \rho'$,*

$$\frac{\partial u(\rho, \theta)}{\partial \theta} \frac{\partial u(\rho', \theta')}{\partial \theta} \leq \frac{\partial u(\rho, \theta')}{\partial \theta} \frac{\partial u(\rho', \theta)}{\partial \theta}. \quad (10)$$

We omit the proof of Corollary 4, which is parallel to that of Corollary 3.

Again, an optimal single-peaked disclosure can be found using optimal control. In particular, when ϕ has a strictly positive density, a unitary single-peaked disclosure signal τ can be represented in the same way as in Section 3.3 but with a reversed matching function $n : [0, \theta^*] \rightarrow [\theta^*, 1]$, where $\theta^* \in (0, 1]$.

4. SPECIAL CASES

We now apply our results to several special cases. These cases are easier to interpret in an equivalent model where the receiver has a private type and makes a binary choice—say, whether to accept or reject a proposal. Specifically, let $G(r|\theta)$ be the conditional distribution of the receiver’s type $r \in [0, 1]$ given the state $\theta \in [0, 1]$. Suppose that $G(r|\theta)$ admits a strictly positive density $g(r|\theta)$ that is continuous in θ . The sender’s and receiver’s utilities from rejection are normalized to zero. The sender’s and receiver’s utilities from acceptance are functions $\tilde{v}(r, \theta)$ and $\tilde{u}(r, \theta)$, such that $\tilde{v}(r, \theta)$ and $\tilde{u}(r, \theta)$ are continuous in θ , $\tilde{v}(r, \theta) > 0$, and $\tilde{u}(r, \theta)g(r|\theta)$ satisfies

¹²More precisely, if the first line of (6) and the second line of (9) hold and in addition at least one of the inequalities in the first line of (6) is strict, then single-peaked disclosure is optimal.

upcrossing in θ and aggregate downcrossing in r . As pointed out by [Kolotilin \(2018, Section 3.2\)](#), this model is equivalent to our model where $\rho \in [0, 1]$ and

$$V(\rho, \theta) = \int_0^\rho \tilde{v}(r, \theta)g(r|\theta)dr \quad \text{and} \quad U(\rho, \theta) = \int_0^\rho \tilde{u}(r, \theta)g(r|\theta)dr.$$

The interpretation is that, before learning r , the receiver chooses a cutoff type ρ such that he accepts if and only if $r \leq \rho$.

4.1. Linear Utilities. The special case where utilities $V(\rho, \theta)$ and $U(\rho, \theta)$ are linear in the state θ has attracted much attention in the persuasion literature (see, for example, [Gentzkow and Kamenica \(2016\)](#), [Kolotilin, Mylovanov, Zapechelnyuk, and Li \(2017\)](#), [Kolotilin \(2018\)](#), [Dworczak and Martini \(2019\)](#)). The key simplification afforded by this case is that the sender's payoff depends only on the mean of the receiver's posterior.

Following [Kolotilin, Mylovanov, Zapechelnyuk, and Li \(2017\)](#), without loss of generality we restrict attention to the case where r is independent of θ , $\tilde{v}(r, \theta) = 1$, and $\tilde{u}(r, \theta) = \theta - r$. So, given a posterior μ , the receiver accepts if and only if the private cost r does not exceed the expected value $\mathbb{E}_\mu[\theta]$, and thus the sender's expected utility is $G(\mathbb{E}_\mu[\theta])$. We call this the *linear case*. Note that this case is canonical.

In the linear case, for $\theta' = p\theta + (1-p)\theta''$ where $p \in (0, 1)$ and $\theta < \theta''$, we have

$$\begin{pmatrix} v(\rho, \theta') \\ u(\rho, \theta') \\ \frac{\partial u(\rho, \theta')}{\partial \rho} \end{pmatrix} = p \begin{pmatrix} v(\rho, \theta) \\ u(\rho, \theta) \\ \frac{\partial u(\rho, \theta)}{\partial \rho} \end{pmatrix} + (1-p) \begin{pmatrix} v(\rho, \theta'') \\ u(\rho, \theta'') \\ \frac{\partial u(\rho, \theta'')}{\partial \rho} \end{pmatrix}.$$

This violates the non-singularity condition (2) in [Theorem 1](#), so it can be generically optimal to pool an interval of states. Since $\rho^*(\mu) = \mathbb{E}_\mu[\theta]$, [Corollary 2](#) yields that full disclosure is optimal if G is convex.

Since $\partial u(\rho, \theta)/\partial \theta = g(\rho)$ is both log-supermodular and log-submodular, [Corollaries 3 and 4](#) yield the following result.

Proposition 1. *In the linear case, both single-dipped disclosure and single-peaked disclosure are optimal.*

[Proposition 1](#) shows that simple disclosure patterns are always optimal in the linear case, regardless of the shape of G ; at the same time, there is generally a wide range of optimal signals. To illustrate this, suppose that $\mu, \mu' \in \text{supp}(\tau)$ induce cutoff types $\rho, \rho' \in [0, 1]$, and suppose that $\theta' = p\theta + (1-p)\theta''$ where $p \in (0, 1)$, $\theta' \in \text{supp}(\mu')$, and $\theta, \theta'' \in \text{supp}(\mu)$. Then a perturbation that shifts weights $\hat{\alpha} = p\hat{\beta}$ and $\hat{\gamma} = (1-p)\hat{\beta}$ on

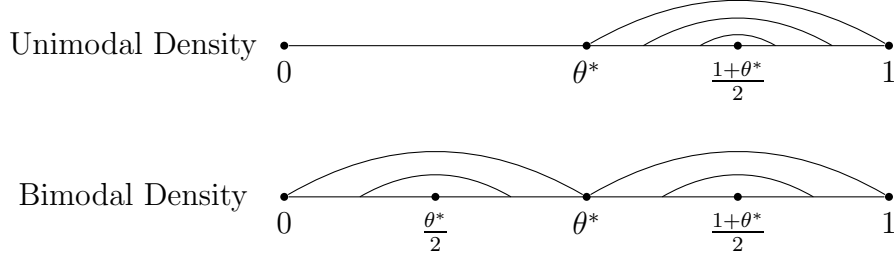


FIGURE 5. Assortative Disclosure for Unimodal and Bimodal Densities

Notes: The top figure shows an optimal assortative signal that is equivalent to separating θ below θ^* and pooling θ above θ^* . The bottom figure shows an optimal assortative signal that is equivalent to pooling θ below θ^* and separately pooling θ above θ^* .

θ and θ'' from μ to μ' and shifts weight $\hat{\beta}$ on θ' from μ' to μ holds fixed ρ and ρ' and their probabilities.¹³

The examples below illustrate various subtleties.

Example 2. By [Kolotilin, Mylovanov, Zapechelnyuk, and Li \(2017, Example 1\)](#), if the density g is unimodal, there is an optimal signal that separates the states below a cutoff θ^* and pools the states above the cutoff. There exists an equivalent signal that is both unitary single-dipped disclosure and unitary single-peaked disclosure. For example, suppose that θ is uniformly distributed on $[0, 1]$. An equivalent unitary assortative signal separates each state $\theta < \theta^*$, so that $\mu = 1\delta_\theta + 0\delta_1$, and pools each pair of states θ and $1 + \theta^* - \theta$ for $\theta \in [\theta^*, (1 + \theta^*)/2]$, so that $\mu = \delta_\theta/2 + \delta_{1 + \theta^* - \theta}/2$. The induced action is θ for $\theta < \theta^*$ and $(1 + \theta^*)/2$ for $\theta \geq \theta^*$, which is both single-dipped and single-peaked on $[0, 1]$.¹⁴ See [Figure 5](#).

Example 3. By [Kolotilin \(2018, Proposition 3, part 2\)](#), for some bimodal g , there is an optimal signal that pools the states below a cutoff θ^* and separately pools the states above θ^* . There exists an equivalent signal that is both single-dipped disclosure and single-peaked disclosure, but there does not exist an equivalent *unitary* assortative signal. For example, suppose that θ is uniformly distributed on $[0, 1]$. An equivalent assortative signal pools each pair of states θ and $\theta^* - \theta$ for $\theta \in (0, \theta^*/2)$ and each pair of states θ and $1 + \theta^* - \theta$ for $\theta \in (\theta^*, (1 + \theta^*)/2)$. See [Figure 5](#).

¹³[Candogan \(2019, Theorem 1\)](#) uses convex programming to prove a related result, under the additional assumption that the receiver has a finite number of actions (or, in the informed-receiver interpretation, a finite number of types).

¹⁴[Terstiege and Wasser \(2019\)](#) study “extensionproof” information disclosure in the linear case. Their [Theorem 1](#) shows that an optimal extensionproof signal takes the form of an assortative signal in [Example 2](#), whereas simply pooling all states above θ^* is suboptimal. In contrast, these two signals are equivalent in our model.

4.2. Translation-Invariant Receiver Utility. Consider the canonical case under the additional assumptions that the receiver's type r is independent of the state θ , and the receiver's utility is translation invariant: $\tilde{u}(r, \theta) = Q(\theta - r)$ for all θ and r and some differentiable, strictly increasing function Q . Let q denote the derivative of Q . We call this the *translation-invariant case*.

An interpretation is that the sender's proposal is a gamble that pays off with probability increasing in $\theta - r$. Suppose the gamble pays $1 - \kappa$ with probability $P(\theta - r)$, and pays $-\kappa$ otherwise, where $\kappa \in (0, 1)$. The receiver's utility is 0 if he rejects the gamble. Then, the receiver's expected utility from accepting is $\tilde{u}(r, \theta) = P(\theta - r) - \kappa$, and the sender's utility from inducing a cutoff type ρ is $G(\rho)$. For example, the gamble may correspond to undertaking a challenging project, where the difficulty of the project is $1 - \theta$, the receiver's ability is $1 - r$, the receiver's bad luck ε has distribution P , and the project succeeds if and only if $1 - \theta \leq 1 - r - \varepsilon$, or equivalently $\varepsilon \leq \theta - r$.

Corollaries 3 and 4 simplify as follows:¹⁵

Proposition 2. *In the translation-invariant case, single-dipped disclosure is optimal if q is log-concave, and single-peaked disclosure is optimal if q is log-convex.*

Intuitively, Corollary 3 says that single-dipped disclosure is optimal if $\log(\partial\tilde{u}(r, \theta)/\partial\theta) = \log q(\theta - r)$ is supermodular in (r, θ) , and $\log q(\theta - r)$ is supermodular in (r, θ) if and only if $\log q$ is concave. For instance, if q is twice continuously differentiable, supermodularity and concavity of $\log q$ both correspond to a negative second derivative.

Proof. If q is log-concave and $t > 0$, then $q(s + t)/q(s)$ decreases in s , and thus for all $\theta < \theta'$ and $\rho < \rho'$ we have

$$\frac{q(\theta' - \rho')}{q(\theta - \rho')} \geq \frac{q(\theta' - \rho)}{q(\theta - \rho)},$$

which is equivalent to (7). Thus, single-dipped disclosure is optimal by Corollary 3.

Similarly, if q is log-convex, then for all $\theta < \theta'$ and $\rho < \rho'$ we have

$$\frac{q(\theta' - \rho')}{q(\theta - \rho')} \leq \frac{q(\theta' - \rho)}{q(\theta - \rho)},$$

and thus single-dipped disclosure is optimal by Corollary 4. \square

Bagnoli and Bergstrom (2005, Tables 1 and 3) report distributions with log-concave and log-convex probability densities. Every log-concave density is single-peaked, whereas every log-convex density is single-dipped. Moreover, most common single-peaked and single-dipped densities are log-concave and log-convex, respectively.

We now consider the limiting case where Q is a step function, so the offered gamble is degenerate. This can be approximated by a sequence of log-concave densities

¹⁵Friedman and Holden (2008, Footnote 11) remark that their informative signal property reduces to the monotone likelihood ratio property under a translation-invariance assumption. This is exactly analogous to the relationship between Corollary 3 and Proposition 2.

increasingly concentrated around a point; so, intuitively, single-dipped disclosure must be optimal.

Formally, the limiting case is given by $\tilde{u}(r, \theta) = \mathbb{1}_{\theta \geq r} - \kappa$. In this case, unitary single-dipped disclosure is optimal. Moreover, whenever the seller discloses that the state is either θ or $\theta' > \theta$, the receiver accepts if and only if $r \leq \theta'$.

Proposition 2'. *Consider the canonical case with $\tilde{u}(r, \theta) = \mathbb{1}_{\theta \geq r} - \kappa$ where $\kappa \in (0, 1)$, and suppose that the prior ϕ has a strictly positive density. There exists an optimal signal τ^* that is unitary single-dipped disclosure and its paired states $\theta < \theta'$ satisfy*

$$\kappa\phi([0, \theta]) = (1 - \kappa)\phi([\theta', 1]).$$

Proof. Consider an arbitrary signal, and let ψ be the probability measure over ρ that this signal induces. Notice that the sender's expected utility is $\int G(\rho)\psi(d\rho)$, where G is increasing. Thus, a higher ψ in the first-order stochastic dominance yields a higher expected utility of the sender. To induce a type- r receiver to accept the proposal, the receiver must believe that the state is at least r with probability at least κ . Thus, by Bayes Rule, $\psi([r, 1]) \leq \min\{1, \phi([r, 1])/\kappa\}$. The signal τ^* achieves this upper bound because the paired states $\{\theta, \theta'\}$ induce the cutoff type $r = \theta'$ to accept, and the probability that τ^* induces at least this type to accept is

$$\psi([r, 1]) = \phi([\theta', 1]) + \phi([0, \theta]) = \phi([\theta', 1]) + \frac{1 - \kappa}{\kappa}\phi([\theta', 1]) = \frac{\phi([r, 1])}{\kappa}.$$

Thus, τ^* is optimal. \square

Another way to show that the signal τ^* is optimal is to notice that it achieves a Pareto-efficient outcome and leaves no rent to the receiver. In particular, the sender would not be able to achieve a higher utility even if she knew the receiver's type r and could condition the signal on it.

The limiting case $\tilde{u}(r, \theta) = \mathbb{1}_{\theta \geq r} - \kappa$ admits two interesting interpretations. First, suppose the sender hopes to sell access to a new technology to the receiver for price κ . The receiver has a privately known status-quo technology r , and assigns value 1 to having the best available technology—for example, the receiver may fear that if $\theta > r$ then a competitor may later acquire technology θ and take over the market. In this case, Proposition 2' shows that the optimal signal pairs states θ and $\theta' > \theta$ such that all receivers with $r \leq \theta'$ accept. Interestingly, the outcome is efficient: for each θ , all receivers with $r \leq \theta$ obtain access to the new technology.

Second, consider the uninformed receiver interpretation, where the receiver chooses action $\rho \in [0, 1]$, and utility is $U(\rho, \theta) = \min\{\rho, \theta\} - \kappa\rho$ for the receiver and $V(\rho)$ for the sender. For an interpretation, suppose the sender is an investment fund manager who can generate excess returns $1 - \kappa$ on each unit of investment up to a capacity of θ (e.g., she has access to an off-market investment opportunity that can profitably absorb θ units of capital), after which she generates negative per-unit excess returns of $-\kappa$. The amount the receiver chooses to invest in the sender's fund is ρ . Here,

Proposition 2' shows that the optimal signal pairs states θ and $\theta' > \theta$ such that the receiver invests $\rho = \theta'$. In particular, the sender's off-market opportunity is always fully exploited, but overinvestment in the sender's fund is so high that the receiver obtains zero expected excess returns.

4.3. Separable Sender Utility. Now assume as in the linear case that r is independent of θ and $\tilde{u}(r, \theta) = \theta - r$, but let the sender's utility take the more general form $\tilde{v}(r, \theta) = w(\theta) > 0$ (independent of r), so that $V(\rho, \theta) = w(\theta)G(\rho)$. This is the most general specification considered by Rayo and Segal (2010, Section VIII).¹⁶ We call this the *separable case*.

In the separable case, Corollary 2' yields that full disclosure is optimal if G is convex and w is increasing, which reproduces Rayo and Segal (2010, Lemma 11). Moreover, the non-singularity condition (2) in Theorem 1 is violated if and only if the first row is a linear combination of the last two rows: that is, $(\theta, w(\theta))$, $(\theta', w(\theta'))$, and $(\theta'', w(\theta''))$ lie on a straight line. Thus, only in this situation can pooling states θ , θ' , and θ'' be optimal. This is a restatement of Rayo and Segal (2010, Lemma 12).

Theorems 3 and 4 simplify as follows:¹⁷

Proposition 3. *In the separable case, single-dipped disclosure is optimal if w is convex, and single-peaked disclosure is optimal if w is concave.*

To prove Proposition 3, suppose that $\mu, \mu' \in \text{supp}(\tau)$ induce cutoff types $\rho, \rho' \in [0, 1]$, and suppose that $\theta' = p\theta + (1 - p)\theta''$ where $p \in (0, 1)$, $\theta' \in \text{supp}(\mu')$, and $\theta, \theta'' \in \text{supp}(\mu)$. Then a perturbation that shifts weights $\hat{\alpha} = p\hat{\beta}$ and $\hat{\gamma} = (1 - p)\hat{\beta}$ on θ and θ'' from μ to μ' and shifts weight $\hat{\beta}$ on θ' from μ' to μ holds fixed ρ and ρ' and changes the sender's expected utility by

$$\hat{\beta}(G(\rho') - G(\rho))(pw(\theta) + (1 - p)w(\theta'') - w(\theta')).$$

Thus, this perturbation increases the sender's expected utility if $\rho' > \rho$ and w is convex, or if $\rho' < \rho$ and w is concave.

Rayo and Segal (2010, Lemma 13) show that an optimal signal τ never induces posteriors $\mu, \mu' \in \text{supp}(\tau)$ such that their pooling segments have an interior intersection point (where the "pooling segment" of a posterior μ is the convex hull of $(\theta, w(\theta))$ for $\theta \in \text{supp}(\mu)$). Figure 6 illustrates the relationship. When w is either convex or concave, this implies that disclosure is assortative: for a convex or concave w , pooling

¹⁶More precisely, Rayo and Segal (2010) assume that the state (ω, θ) is two-dimensional, where ω is the sender's value of the proposal so that $\tilde{v}(r, \theta, \omega) = \omega$, and θ is the receiver's value of the proposal so that $\tilde{u}(r, \theta) = \theta - r$. They assume that there are finitely many states (θ, ω) , so generically (that is, unless there exist distinct states (θ, ω) that have the same component θ) we can write the sender's utility as $\tilde{v}(r, \theta) = w(\theta)$.

¹⁷Proposition 3 can be established by checking the conditions of Theorems 3 and 4, but it is easier to apply Lemmas 3 and 4 directly. The same comment applies to Proposition 5 in Section 4.5.

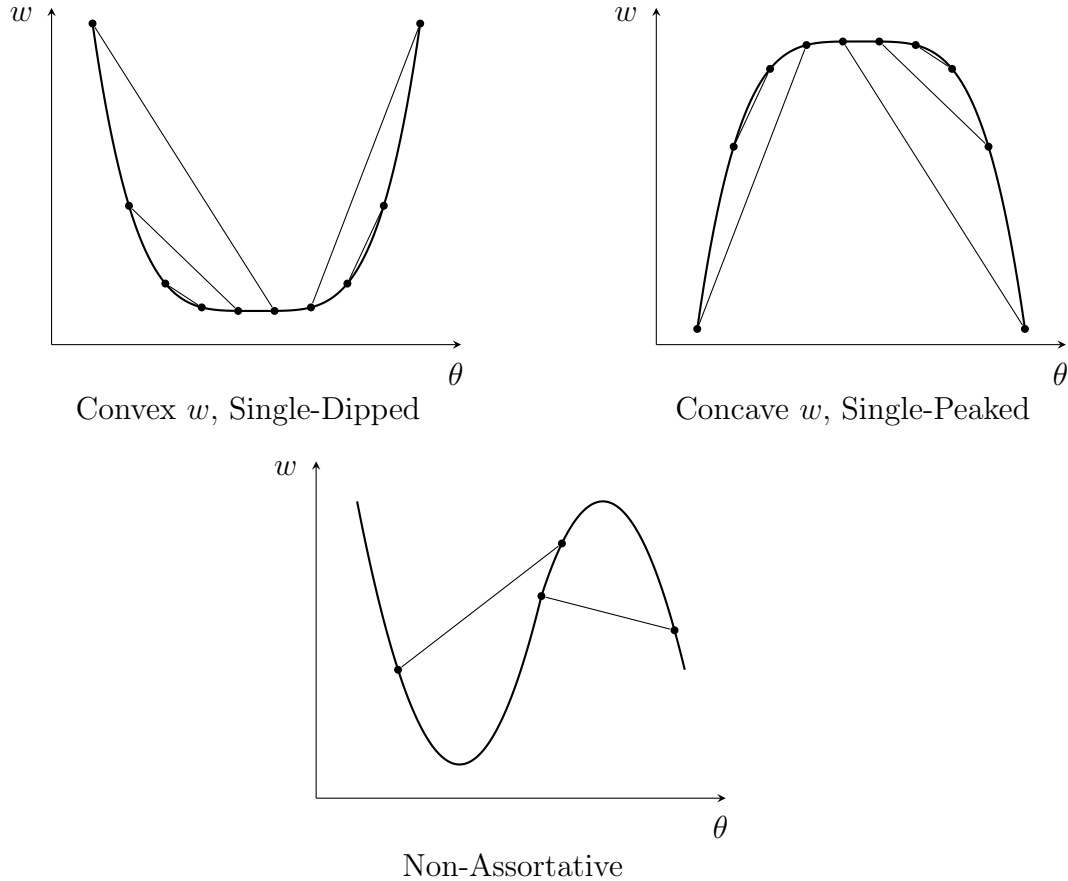


FIGURE 6. Convex/Concave w and Single-Dipped/Single-Peaked τ

Notes: When w is either convex or concave, pooling segments of assortative disclosure do not intersect, and, conversely, a signal with non-intersecting pooling segments is assortative. When w is neither convex nor concave, a signal with non-intersecting pooling segments can be non-assortative.

segments of assortative disclosure do not intersect, and, conversely, a signal with non-intersecting pooling segments is assortative. However, [Rayo and Segal](#)'s analysis does not establish that, in addition to being assortative, optimal disclosure is single-dipped when w is convex and single-peaked when w is concave.¹⁸

¹⁸[Rayo and Segal](#) remark that “[b]eyond these results [Lemmas 11–13], little can be said about the optimal pooling graph for arbitrary G , given that its curvature can greatly influence the outcome.” Our results on the shape of the “optimal pooling graph” coincide with [Rayo and Segal](#)'s, while Proposition 3 provides additional information on how the receiver's optimal action varies within this graph.

Now consider the separable case with the additional assumption that G is uniform. This is the main specification considered by [Rayo and Segal \(2010\)](#), as well as by [Nikandrova and Pans \(2017\)](#) in their study of information disclosure in sequential auctions. Under this extra assumption, we can show that if w is convex or concave, then *unitary* assortative disclosure is optimal. This is because, when G is uniform, an optimal signal must satisfy the following two properties, implied by condition (3), and also by [Rayo and Segal \(2010, Lemmas 1 and 2\)](#). First, non-degenerate paired states $\{\theta, \theta'\}$ have a negative slope in the sense that

$$\frac{w(\theta') - w(\theta)}{\theta' - \theta} \leq 0;$$

otherwise it would be better to separate these states. Second, the means of a pair of paired states have a positive slope in the sense that

$$\frac{(p_2 w(\theta_2) + (1 - p_2) w(\theta'_2)) - (p_1 w(\theta_1) + (1 - p_1) w(\theta'_1))}{(p_2 \theta_2 + (1 - p_2) \theta'_2) - (p_1 \theta_1 + (1 - p_1) \theta'_1)} \geq 0;$$

otherwise it would be better to pool these pairs. It can be seen graphically from [Figure 7](#) that, when w is convex, single-dipped disclosure that satisfies these two properties is unitary. This corresponds to the main result in [Nikandrova and Pans \(2017, Lemma 2 and Theorem 3\)](#).

To conclude, we provide an example in which assortative disclosure is suboptimal. Let r and θ be independently uniformly distributed on $[0, 1]$, $\tilde{u}(r, \theta) = \theta - r$, and

$$\tilde{v}(r, \theta) = \begin{cases} \theta + 1/2, & \text{if } \theta \in [0, 1/2], \\ \theta - 1/2, & \text{if } \theta \in (1/2, 1]. \end{cases}$$

By [Dworczak and Kolotilin \(2019, Proposition 2\)](#), there is an optimal signal that pools each pair of states θ and $1/2 + \theta$ for $\theta \in (0, 1/2)$. This disclosure pattern corresponds to the “median matching” of [Kremer and Maskin \(1996\)](#) and [Legros and Newman \(2002\)](#), in which matching is “positively assortative across the median.”

4.4. Overvaluation Preferences. Now assume that the receiver’s type r is uniformly distributed on $[0, 1]$ independently of θ , and the sender’s and receiver’s utilities are $\tilde{u}(r, \theta) = \theta - r$ and $\tilde{v}(r, \theta) = (\max\{r - \theta, 0\})^\zeta$ with $\zeta \geq 1$, so that $V(\rho, \theta) = (\max\{\rho - \theta, 0\})^{\zeta+1}/(\zeta + 1)$.

The interpretation is that the sender wants the receiver to overvalue the state, and does not care about undervaluation; we thus call this the *overvaluation case*.

For a possible application, suppose that the sender is a school, the proposal is a graduating student, and the receiver is a prospective employer. The state θ is the student’s ability, and the type r is the employer’s quality. The employer hires the student if the student’s expected ability is above the employer’s quality. The school, however, wants to overplace the student and does not care about underplacement (perhaps because it can advertise overplacements to prospective future students and

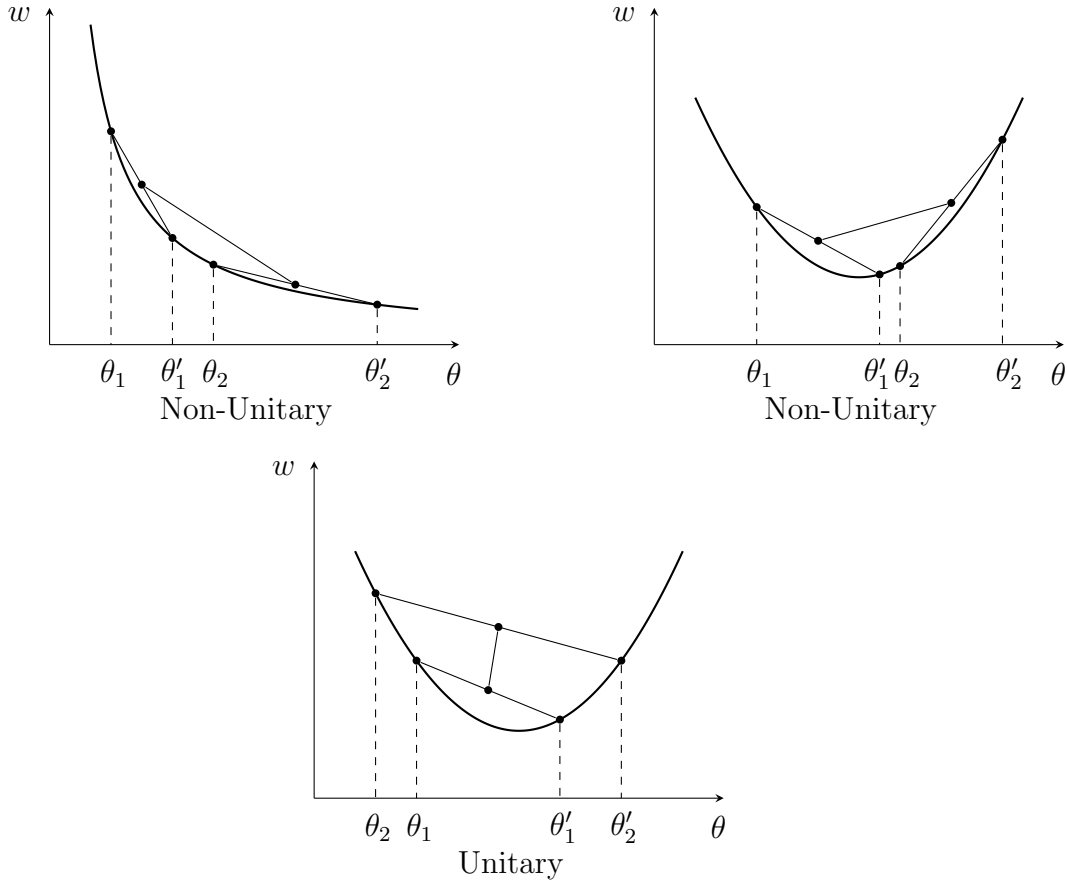


FIGURE 7. Convex w , Uniform G , and Unitary Assortative τ

Notes: A non-unitary assortative signal must have ordered pairs $\theta_1 < \theta'_1 \leq \theta_2 < \theta'_2$. The top-left figure shows that if the ordered pairs $\{\theta_1, \theta'_1\}$ and $\{\theta_2, \theta'_2\}$ both have negative slopes, then their means have a negative slope as well. The top-right figure shows that if the means of the ordered pairs $\{\theta_1, \theta'_1\}$ and $\{\theta_2, \theta'_2\}$ have a positive slope, then the paired states $\{\theta_2, \theta'_2\}$ have a positive slope as well. A unitary assortative signal has only nested pairs $\theta_2 \leq \theta_1 \leq \theta'_1 \leq \theta'_2$. The bottom figure shows that the nested pairs $\{\theta_1, \theta'_1\}$ and $\{\theta_2, \theta'_2\}$ can have negative slopes while their means have a positive slope.

donors, while concealing underplacements). The school's utility is an increasing convex function of the degree of overplacement, $\max\{r - \theta, 0\}$ (where convexity may reflect increasing returns to truly spectacular overplacements).

With these preferences, $V(\rho, \theta)$ is not strictly increasing in ρ and thus it violates Assumption 2; so we cannot apply Lemma 3 directly. Indeed, A is not well-defined

because the denominator is zero. When $V(\rho, \theta)$ is (weakly) increasing in ρ , the lemma takes the following form (with the same proof):

Lemma 3’. *Single-dipped disclosure is optimal if, for all $\theta < \theta' < \theta''$ and $\rho < \rho'$ such that $u(\rho, \theta) < 0 < u(\rho, \theta'')$, there exists a non-zero vector $(\hat{\alpha}, \hat{\beta}, \hat{\gamma}) \geq 0$ such that*

$$\begin{pmatrix} u(\rho', \theta) & -u(\rho', \theta') & u(\rho', \theta'') \\ -u(\rho, \theta) & u(\rho, \theta') & -u(\rho, \theta'') \\ V(\rho', \theta) - V(\rho, \theta) & -(V(\rho', \theta') - V(\rho, \theta')) & V(\rho', \theta'') - V(\rho, \theta'') \end{pmatrix} \begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \\ \hat{\gamma} \end{pmatrix} \geq \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \quad (11)$$

Lemma 3’ yields the following result.

Proposition 4. *In the overvaluation case, single-dipped disclosure is optimal.*

For intuition, suppose that $\mu, \mu' \in \text{supp}(\tau)$ induce cutoff types $\rho < \rho'$, and $\theta' \in \text{supp}(\mu')$ lies between $\theta, \theta'' \in \text{supp}(\mu)$. We focus on the case $\theta' < \rho'$. Consider a perturbation that shifts weight $\hat{\alpha}$ on θ from μ to μ' and shifts weight $\hat{\beta}$ on θ' from μ' to μ in such a way that ρ' holds fixed. Note that $\hat{\alpha}/\hat{\beta} = (\rho' - \theta') / (\rho' - \theta)$, while the corresponding ratio to hold ρ fixed would be $(\rho - \theta') / (\rho - \theta)$, which is smaller; hence, this perturbation increases ρ .¹⁹ The perturbation also increases the sender’s expected utility for fixed ρ and ρ' , intuitively because it negatively rearranges probability mass across actions, which increases the sender’s expected utility, given that it is sufficiently submodular.

A potentially very important application where the sender has overvaluation-like preferences is the design of bank stress tests. Goldstein and Leitner (2018) consider a model of stress test design where the regulator (sender) wants to prevent bank runs by keeping banks’ cash holdings above a threshold θ_0 . This can be achieved by designing a test that pools distressed assets with true values below θ_0 with higher-value assets, so that the resulting bundle can be sold at a price above θ_0 . Goldstein and Leitner show that unitary single-dipped disclosure is optimal, using the same perturbation as that underlying Proposition 4. While their sender preferences do not exactly fall into our overvaluation class, in Appendix A.3 we show how to nest their model in ours, and we reproduce their result using a slight modification of the proof of Proposition 4.

4.5. Affiliated Information. In all the special cases considered so far, the receiver’s private type r directly enters his utility and is independent of θ , and is thus naturally interpreted as a private taste shock. We next turn to the case studied by Guo and Shmaya (2019), where the density $g(r|\theta)$ is log-submodular, $\tilde{v}(r, \theta) = 1$, and $\tilde{u}(r, \theta) = \theta - \theta_0$ for $\theta_0 \in (0, 1)$. Here the interpretation is that it is commonly known that the receiver wishes to accept if and only if the prospect’s value θ exceeds a threshold θ_0 ,

¹⁹Here is a physical intuition for this result: Consider two balanced levers, with fulcrums located at ρ and ρ' . Suppose weight located at θ is moved from the first lever to the second, while weight located at θ' is moved from the second lever to the first, so as to keep the second lever balanced. Then the first lever tips up.

and the receiver's type r is his private signal of θ , where a high r indicates that θ is more likely to be low.²⁰ We call this the *affiliated information case*.

In this case, $V(\rho, \theta) = G(\rho|\theta)$ and $u(\rho, \theta) = (\theta - \theta_0)g(\rho|\theta)$. These preferences satisfy Assumptions 1–3. In particular, $u(\rho, \theta)$ satisfies aggregate downcrossing in ρ , by Theorem A1 and Remark A1.

Theorem 4 now yields the following result, which reproduces the main qualitative insight of Guo and Shmaya (2019, Theorem 3.1).

Proposition 5. *In the affiliated information case, unitary single-peaked disclosure is optimal.*

To prove Proposition 5, fix $\theta < \theta' < \theta''$ and $\rho < \rho'$ such that $\theta < \theta_0 < \theta''$, and suppose there exist posteriors $\mu, \mu' \in \text{supp}(\tau)$ inducing cutoff types $\rho, \rho' \in [0, 1]$ with $\theta' \in \text{supp}(\mu)$, $\theta, \theta'' \in \text{supp}(\mu')$. We focus on the case $\theta_0 < \theta'$. Consider a perturbation that shifts weight $\hat{\beta}$ on θ' from μ to μ' and shifts weight $\hat{\gamma}$ on θ'' from μ' to μ where

$$\hat{\beta} = \frac{\varepsilon}{(\theta' - \theta_0)g(\rho|\theta')} \quad \text{and} \quad \hat{\gamma} = \frac{\varepsilon}{(\theta'' - \theta_0)g(\rho|\theta'')} \quad \text{for } \varepsilon > 0.$$

This perturbation holds ρ fixed and increases ρ' : by log-submodularity of g ,

$$u(\rho', \theta')\hat{\beta} - u(\rho', \theta'')\hat{\gamma} = \varepsilon \left(\frac{g(\rho'|\theta')}{g(\rho|\theta')} - \frac{g(\rho'|\theta'')}{g(\rho|\theta'')} \right) \geq 0.$$

The intuition is that, since a type- ρ receiver is more optimistic about θ than a type- ρ' receiver, he assigns higher prior probability to θ'' relative to θ' . He therefore finds a signal that puts more weight on θ'' relatively more persuasive, while the more pessimistic type- ρ' receiver is more persuaded by a signal that puts more weight on θ' .

Moreover, this perturbation increases the sender's expected utility for fixed ρ and ρ' :

$$\begin{aligned} & (V(\rho', \theta') - V(\rho, \theta'))\hat{\beta} - (V(\rho', \theta'') - V(\rho, \theta''))\hat{\gamma} \\ &= \varepsilon \left(\frac{G(\rho'|\theta') - G(\rho|\theta')}{(\theta' - \theta_0)g(\rho|\theta')} - \frac{G(\rho'|\theta'') - G(\rho|\theta'')}{(\theta'' - \theta_0)g(\rho|\theta'')} \right) \\ &> \frac{\varepsilon}{(\theta' - \theta_0)} \left(\frac{G(\rho'|\theta') - G(\rho|\theta')}{g(\rho|\theta')} - \frac{G(\rho'|\theta'') - G(\rho|\theta'')}{g(\rho|\theta'')} \right) \\ &= \frac{\varepsilon}{(\theta' - \theta_0)} \int_{\rho}^{\rho'} \left(\frac{g(r|\theta')}{g(\rho|\theta')} - \frac{g(r|\theta'')}{g(\rho|\theta'')} \right) dr \geq 0, \end{aligned}$$

where the first inequality is by $\theta_0 < \theta' < \theta''$ and the second inequality is by log-submodularity of g . Thus, single-peaked disclosure is optimal.²¹

²⁰This ordering convention on r is opposite to Guo and Shmaya's. It is dictated by our convention that the receiver accepts for types *below* the cutoff.

²¹The proof in the case $\theta < \theta' \leq \theta_0$ is similar, except that the relevant perturbation holds ρ' fixed and increases ρ .

The above chain of inequalities can be given an intuitive explanation. There are two effects that both benefit the sender. First, if we ignore the effect of θ on $g(\rho|\theta)$ (for example, suppose for the moment that $g(\rho|\theta') = g(\rho|\theta'')$), then the fact that \tilde{u} is strictly increasing in θ implies that $\hat{\beta} > \hat{\gamma}$. That is, to keep a type- ρ receiver indifferent, the weight on θ' that can be moved from μ to μ' is greater than the weight on θ'' that moves in the opposite direction. This benefits the sender, as more weight moves from the lower receiver cutoff to the higher cutoff than moves in the opposite direction. This explains the first inequality.

The second inequality comes from the fact that, since r is more likely to be high when θ is low, $\Pr(r \in [\rho, \rho']|\theta)/\Pr(r = \rho|\theta)$ is decreasing in θ . That is, the higher is θ , the lower is the relative probability that an increase in the receiver’s cutoff from ρ to ρ' is pivotal for the receiver’s action. Thus, moving mass on θ' to μ' while moving mass on θ'' to μ benefits the sender by inducing a higher receiver cutoff at those states where the choice of cutoff is more likely to matter for the receiver’s action, and inducing a lower cutoff at states where the choice of cutoff is less likely to matter. This explains the second inequality.

Finally, it is easy to see that *unitary* single-peaked disclosure is optimal. Otherwise, there are paired states on the same side of θ_0 , and the sender’s expected utility does not change if these states are separated.²²

5. CONCLUSION

This paper has characterized the optimality of simple patterns of information disclosure. We provided general conditions for the optimality of full disclosure, as well as—in our main results—the optimality of single-dipped and single-peaked disclosure, where the induced receiver action is single-dipped or single-peaked on each set of nested pairs of states. Most of our results are driven by relatively simple complementarity-type conditions, especially log-super/submodularity of the cross-partial derivative of the receiver’s utility. Our results generalize a wide range of previous findings and generate new applications, for example to inducing investment in a project with an unknown capacity to absorb capital.

Several directions remain to be explored. First, one could try to characterize more complex disclosure patterns, such as the median matching discussed in Section 4.3 as well as in Kremer and Maskin (1996) and Legros and Newman (2002). More progress could also be made on explicitly solving for optimal single-dipped/single-peaked signals, as in the optimal control approach discussed in Sections 3.3 and 3.4. Alternatively, one could investigate when simpler disclosure patterns are exactly or approximately optimal. In particular, the “negative assortative” matching patterns characterized in the present paper are already fairly complicated, and in some settings the sender may be effectively constrained to use simpler disclosure schemes, such as

²²The above proof can be easily extended to show that unitary single-peaked disclosure remains optimal in the more general setting considered in Guo and Shmaya (2019, Section 5.1).

partitioning states into intervals. (For instance, in reality it is probably not feasible to design a stress test that pools only the weakest and strongest banks.) Finally, while we have sketched several applications that seem promising to us, the details are left to future work. One application to which our approach is particularly well suited is the problem of gerrymandering, or designing optimal electoral districts. We explore this application in a companion paper.

APPENDIX

A.1. Characterization of Aggregate Downcrossing. We now present necessary and sufficient conditions for aggregate downcrossing, which are established by [Quah and Strulovici \(2012, Theorem 1\)](#) and [Choi and Smith \(2017, Corollary 2\)](#). Let u_+ and u_- be the positive and negative parts of u , so that $u_+(\rho, \theta) = \max\{0, u(\rho, \theta)\}$ and $u_-(\rho, \theta) = -\min\{0, u(\rho, \theta)\}$. Notice that $u_+(\rho, \theta) \geq 0$, $u_-(\rho, \theta) \geq 0$, and $u(\rho, \theta) = u_+(\rho, \theta) - u_-(\rho, \theta)$.

Theorem A1. *Function $u(\rho, \theta)$ satisfies aggregate downcrossing in ρ if and only if*

$$(i) \ u(\rho, \theta) \text{ satisfies downcrossing in } \rho, \text{ so that for all } \theta \text{ and } \rho < \rho', \quad (12)$$

$$u(\rho', \theta) \geq (>)0 \implies u(\rho, \theta) \geq (>)0, \text{ and}$$

(ii) $u(\rho, \theta)$ satisfies signed-ratio monotonicity in ρ , so that for all θ, θ' , and $\rho < \rho'$,

$$u_-(\rho', \theta)u_+(\rho, \theta') \geq u_-(\rho, \theta)u_+(\rho', \theta'). \quad (13)$$

Proof. If. This proof is an adaptation of that of [Karlin and Rubin \(1956, Lemma 1\)](#). Let $\mathcal{U}_+(\rho) = \int u_+(\rho, \theta)d\mu$ and $\mathcal{U}_-(\rho) = \int u_-(\rho, \theta)d\mu$, so that $\mathcal{U}_+(\rho) \geq 0$, $\mathcal{U}_-(\rho) \geq 0$, and $\mathcal{U}(\rho) = \int u(\rho, \theta)d\mu = \mathcal{U}_+(\rho) - \mathcal{U}_-(\rho)$.

Since $u(\rho, \theta)$ satisfies signed-ratio monotonicity, we have, for $\rho < \rho'$,

$$\begin{aligned} & \mathcal{U}_-(\rho')\mathcal{U}_+(\rho) - \mathcal{U}_-(\rho)\mathcal{U}_+(\rho') \\ &= \int \int [u_-(\rho', \theta)u_+(\rho, \theta') - u_-(\rho, \theta)u_+(\rho', \theta')]d\mu(\theta)d\mu(\theta') \geq 0. \end{aligned} \quad (14)$$

Since $u(\rho, \theta)$ satisfies downcrossing in ρ , we have, for $\rho < \rho'$,

$$\begin{aligned} \mathcal{U}_+(\rho') > 0 &\implies \mathcal{U}_+(\rho) > 0, \\ \mathcal{U}_-(\rho) > 0 &\implies \mathcal{U}_-(\rho') > 0. \end{aligned} \quad (15)$$

Combining (14) and (15) yields that, for $\rho < \rho'$,

$$\mathcal{U}_+(\rho') \geq (>)\mathcal{U}_-(\rho') \implies \mathcal{U}_+(\rho) \geq (>)\mathcal{U}_-(\rho).$$

Only if. Since $u(\rho, \theta)$ satisfies aggregate downcrossing in ρ , it satisfies downcrossing in ρ (consider μ that assigns probability one to θ). Suppose, for a contradiction, that $u(\rho, \theta)$ does not satisfy signed-ratio monotonicity in ρ , so that there exist θ, θ' , and $\rho < \rho'$ such that

$$u_-(\rho', \theta)u_+(\rho, \theta') < u_-(\rho, \theta)u_+(\rho', \theta').$$

Since $u_- \geq 0$ and $u_+ \geq 0$, we have $u_-(\rho, \theta) > 0$ and $u_+(\rho', \theta') > 0$. Thus, $u_-(\rho', \theta) > 0$ and $u_+(\rho, \theta') > 0$ because $u(\rho, \theta)$ satisfies downcrossing in ρ . Consider μ given by $\mu(\theta) = p$ and $\mu(\theta') = 1 - p$ where

$$p = \frac{u_+(\rho', \theta')}{u_-(\rho', \theta) + u_+(\rho', \theta')}.$$

Clearly, $\mathcal{U}(\rho') = 0$ but

$$\mathcal{U}(\rho) = \frac{-u_+(\rho', \theta')u_-(\rho, \theta) + u_-(\rho', \theta)u_+(\rho, \theta')}{u_-(\rho', \theta) + u_+(\rho', \theta')} < 0,$$

which leads to a contradiction that $u(\rho, \theta)$ satisfies aggregate downcrossing in ρ . \square

Remark A1. Let (12) hold. Then (13) holds if and only if for all θ, θ' , and $\rho < \rho'$ such that $u(\rho, \theta) < 0 < u(\rho, \theta')$ and $u(\rho', \theta) < 0 < u(\rho', \theta')$, we have

$$-\frac{u(\rho', \theta)}{u(\rho', \theta')} \geq -\frac{u(\rho, \theta)}{u(\rho, \theta')}. \quad (16)$$

Proof. Clearly, (13) holds if either $u(\rho, \theta) \geq 0$ or $u(\rho', \theta') \leq 0$. By (13), $u(\rho', \theta) \geq 0$ implies $u(\rho, \theta) \geq 0$, so (13) holds with equality. Similarly, by (13), $u(\rho, \theta') \leq 0$ implies $u(\rho', \theta') \leq 0$, so (13) holds with equality again. Thus, (13) imposes an additional requirement on u only for θ, θ' , and $\rho < \rho'$ such that $u(\rho, \theta) < 0 < u(\rho, \theta')$ and $u(\rho', \theta) < 0 < u(\rho', \theta')$. This requirement can be written as (16). \square

A.2. Omitted Proofs.

Proof of Lemma 1. Define $\widehat{V}(\mu) = \int V(\rho^*(\mu), \theta) d\mu$. Since U and V are continuous functions defined on the compact set, $\widehat{V}(\mu)$ is upper-semicontinuous in μ , by the Berge Maximum Theorem, and there exists an optimal signal τ^* (see Kamenica and Gentzkow (2011, Section 3 in the Online Appendix) and Dworzak and Kolotilin (2019, Appendix B.3)). Recall that an optimal signal $\tau^* \in \Delta(\Delta([0, 1]))$

$$\begin{aligned} & \text{maximizes } \int_{\Delta([0,1])} \widehat{V}(\mu) d\tau \\ & \text{subject to } \int_{\Delta([0,1])} \mu d\tau = \phi. \end{aligned} \quad (17)$$

Consider first the case where the prior ϕ has finite support. Then we can redefine the state space as $\Theta = \text{supp}(\phi)$. Let $|\Theta|$ denote the number of elements in the set Θ . By Kamenica and Gentzkow (2011, Section 4 in the Online Appendix), there exists an optimal signal τ^* such that $|\text{supp}(\tau^*)| \leq |\Theta|$. This is so, because the problem (17) has an extreme optimal solution τ^* by the Bauer Maximum Principle, and each extreme solution τ satisfies $|\text{supp}(\tau)| \leq |\Theta|$ because (17) has $|\Theta|$ equation constraints.

A belief $\mu \in \text{supp}(\tau^*)$ induces action $\rho^*(\mu)$ given by

$$\rho^*(\mu) = \max \left\{ \rho : \sum_{\theta \in \Theta} u(\rho, \theta) \mu(\theta) \geq 0 \right\}.$$

Consider a nonempty bounded polyhedron $P^\mu \subset \mathbb{R}^{|\Theta|}$ defined by the constraints

$$\begin{aligned} \sum_{\theta \in \Theta} u(\rho^*(\mu), \theta) \eta(\theta) &= \sum_{\theta \in \Theta} u(\rho^*(\mu), \theta) \mu(\theta), \\ \sum_{\theta \in \Theta} \eta(\theta) &= \sum_{\theta \in \Theta} \mu(\theta) = 1, \\ \eta(\theta) &\geq 0, \forall \theta \in \Theta. \end{aligned}$$

Since the polyhedron P^μ is in standard form with only two equation constraints, each extreme point $\eta^\mu \in P^\mu$ satisfies $|\text{supp}(\eta^\mu)| \leq 2$. Moreover, since the polyhedron has a finite number of constraints, there is a finite number of extreme points of P^μ which we denote by $E^\mu = \{\eta_1^\mu, \dots, \eta_{|E^\mu|}^\mu\}$. By the Krein-Milman Theorem, P^μ is the convex hull of E^μ ; so there exist $\{(\lambda_1^\mu, \dots, \lambda_{|E^\mu|}^\mu : \sum_{i=1}^{|E^\mu|} \lambda_i^\mu = 1, \lambda_i^\mu \geq 0)\}$ that represent μ as follows

$$\mu = \sum_{i=1}^{|E^\mu|} \lambda_i^\mu \eta_i^\mu.$$

Since $\eta_i^\mu \in P^\mu$, it follows that $\rho^*(\eta_i^\mu) \geq \rho^*(\mu)$. Since $V(\rho, \theta)$ is strictly increasing in ρ , a signal τ^{**} given by

$$\tau^{**} = \sum_{\mu \in \text{supp}(\tau^*)} \sum_{i=1}^{|E^\mu|} \tau(\mu) \lambda_i^\mu \delta_{\eta_i^\mu}$$

yields a (weakly) higher sender's expected utility than τ^* does, and thus τ^{**} is optimal.

Consider now the case where $\text{supp}(\phi)$ is not a finite set. Let τ^* be an optimal signal. By the Choquet Theorem, for each $\mu \in \text{supp}(\tau^*)$, there exists a probability measure λ^μ that represents μ and is supported on the extreme beliefs E^μ where each extreme belief $\eta^\mu \in E^\mu$ satisfies $\text{supp}(\eta^\mu) \leq 2$ and $\rho^*(\eta^\mu) \geq \rho^*(\mu)$. Next, by the Measurable Selection Theorem, we can construct a signal τ^{**} that is supported on $\cup_{\mu \in \text{supp}(\tau^*)} E^\mu$ and yields a (weakly) higher sender's expected utility than τ^* does. \square

Proof of Theorem 1. Suppose for contradiction that such an optimal signal τ exists. Since ϕ has finite support, we can redefine the state space as $\Theta = \text{supp}(\phi)$. Denote

$$\begin{aligned} (v, v', v'') &= (v(\rho^*(\mu), \theta), v(\rho^*(\mu), \theta'), v(\rho^*(\mu), \theta'')), \\ (u, u', u'') &= (u(\rho^*(\mu), \theta), u(\rho^*(\mu), \theta'), u(\rho^*(\mu), \theta'')), \\ (u_\rho, u'_\rho, u''_\rho) &= \left(\frac{\partial u(\rho^*(\mu), \theta)}{\partial \rho}, \frac{\partial u(\rho^*(\mu), \theta')}{\partial \rho}, \frac{\partial u(\rho^*(\mu), \theta'')}{\partial \rho} \right). \end{aligned}$$

Let us split μ into two identical and equiprobable posteriors μ_1 and μ_2 and then move a sufficiently small probability mass $(\varepsilon, \varepsilon', \varepsilon'') \neq (0, 0, 0)$ at $(\theta, \theta', \theta'')$ from μ_1 to μ_2 , so that posteriors μ_1 and μ_2 are well-defined and satisfy $\text{supp}(\mu_1) = \text{supp}(\mu_2)$ and $\mu_1 \neq \mu_2$. Moreover, $\tau(\mu_1)\mu_1(\theta) = \tau(\mu)\mu(\theta)/2 - \varepsilon$, $\tau(\mu_2)\mu_2(\theta) = \tau(\mu)\mu(\theta)/2 + \varepsilon$, and similar conditions hold for other states in Θ . Suppose further that $(\varepsilon, \varepsilon', \varepsilon'')$ satisfies

$$u\varepsilon + u'\varepsilon' + u''\varepsilon'' = 0, \quad (18)$$

$$u_\rho\varepsilon + u'_\rho\varepsilon' + u''_\rho\varepsilon'' = 0. \quad (19)$$

Since a homogenous system of linear equations with a non-singular matrix has only the zero solution, it follows from $(\varepsilon, \varepsilon', \varepsilon'') \neq (0, 0, 0)$, (2), (18), and (19) that

$$v\varepsilon + v'\varepsilon' + v''\varepsilon'' \neq 0. \quad (20)$$

Since $\text{supp}(\mu_1) = \text{supp}(\mu_2)$ and (18) holds, it follows that $\rho^*(\mu_1) = \rho^*(\mu_2) = \rho^*(\mu)$. Consider now a perturbation that moves an infinitesimal probability mass (dp, dp', dp'') at $(\theta, \theta', \theta'')$ from μ_1 to μ_2 , where (dp, dp', dp'') satisfies

$$udp + u'dp' + u''dp'' \neq 0. \quad (21)$$

Posterior μ_1 induces action $\rho^*(\mu) + d\rho_1 + o(d\rho_1)$ where $d\rho_1$ is given by

$$d\rho_1 = \frac{1}{\tau(\mu_1)} \frac{udp + u'dp' + u''dp''}{\sum_{\theta \in \Theta} \frac{\partial u(\rho^*(\mu_1), \theta)}{\partial \rho} \mu_1(\theta)} = \frac{2}{\tau(\mu)} \frac{udp + u'dp' + u''dp''}{\sum_{\theta \in \Theta} \frac{\partial u(\rho^*(\mu), \theta)}{\partial \rho} \mu(\theta)} \neq 0,$$

where the first equality holds by the Implicit Function Theorem, the second equality holds by (19) and definition of μ_1 , and the non-equality holds by (21). Similarly, μ_2 induces action $\rho^*(\mu) + d\rho_2 + o(d\rho_2)$ where $d\rho_2$ is given by

$$d\rho_2 = -\frac{1}{\tau(\mu_2)} \frac{udp + u'dp' + u''dp''}{\sum_{\theta \in \Theta} \frac{\partial u(\rho^*(\mu_2), \theta)}{\partial \rho} \mu_2(\theta)} = -\frac{2}{\tau(\mu)} \frac{udp + u'dp' + u''dp''}{\sum_{\theta \in \Theta} \frac{\partial u(\rho^*(\mu), \theta)}{\partial \rho} \mu(\theta)} \neq 0.$$

Notice that $d\rho_2 = -d\rho_1$. The perturbation changes the sender's expected utility by $dV + o(d\rho_2)$ where dV is given by

$$\begin{aligned} dV &= d\rho_1 \tau(\mu_1) \sum_{\theta \in \Theta} v(\rho^*(\mu_1), \theta) \mu_1(\theta) + d\rho_2 \tau(\mu_2) \sum_{\theta \in \Theta} v(\rho^*(\mu_2), \theta) \mu_2(\theta) \\ &= d\rho_1 \frac{\tau(\mu)}{2} \sum_{\theta \in \Theta} v(\rho^*(\mu), \theta) \mu(\theta) - d\rho_1 (v\varepsilon + v'\varepsilon' + v''\varepsilon'') \\ &\quad + d\rho_2 \frac{\tau(\mu)}{2} \sum_{\theta \in \Theta} v(\rho^*(\mu), \theta) \mu(\theta) + d\rho_2 (v\varepsilon + v'\varepsilon' + v''\varepsilon'') \\ &= 2d\rho_2 (v\varepsilon + v'\varepsilon' + v''\varepsilon''), \end{aligned}$$

where the first equality holds by the Intermediate Value Theorem, the second equality holds by $\rho^*(\mu_1) = \rho^*(\mu_2) = \rho^*(\mu)$ and definitions of μ_1 and μ_2 , and the third equality holds by $d\rho_2 = -d\rho_1$. Since (20) holds, it follows that $dV \neq 0$. Thus, either this perturbation or the reverse perturbation, where (dp, dp', dp'') is replaced with

$-(dp, dp', dp'')$, strictly increases the sender's expected utility. But this contradicts that τ is optimal. \square

Proof of Theorem 2. If. This is proved in the text.

Only if. This is proved in the text for the case where $\phi(\theta) > 0$ and $\phi(\theta') > 0$, but it holds for the general case where $\theta, \theta' \in \text{supp}(\phi)$, so that $\phi((\theta - r, \theta + r)) > 0$ and $\phi((\theta' - r, \theta' + r)) > 0$ for all $r > 0$. To illustrate, consider the case where ϕ has a strictly positive density f in neighborhoods of θ and θ' . Suppose that (3) is violated for some $\mu = p\delta_\theta + (1 - p)\delta_{\theta'}$, so that

$$pV(\rho^*(\delta_\theta), \theta) + (1 - p)V(\rho^*(\delta_{\theta'}), \theta') < pV(\rho^*(\mu), \theta) + (1 - p)V(\rho^*(\mu), \theta').$$

Clearly, this inequality implies that $p \in (0, 1)$ and $\theta \neq \theta'$; so let $\theta < \theta'$. Consider first the trivial case where $\rho^*(\delta_\theta) < \rho^*(\mu) = \rho^*(\delta_{\theta'}) = 1$. Since $u(\rho, \theta)$ is continuous in θ , a posterior belief that pools some mass of states near θ' with a much smaller mass of states near θ induces action 1 and thus is strictly preferred by the sender to full disclosure of these states.

In the remaining case, which satisfies the inequality, by monotonicity of $V(\rho, \theta)$ in ρ , and downcrossing of $u(\rho, \theta)$ in ρ , we have

$$\begin{aligned} \rho^*(\delta_\theta) &< \rho^*(\mu) < \rho^*(\delta_{\theta'}), \\ u(\rho^*(\mu), \theta) &< 0 < u(\rho^*(\mu), \theta'), \\ pu(\rho^*(\mu), \theta) + (1 - p)u(\rho^*(\mu), \theta') &= 0. \end{aligned}$$

Since $u(\rho, \theta)$ is continuous in θ , there exist $\varepsilon > 0$ and $\varepsilon' > 0$ such that

$$\int_{\theta - \varepsilon}^{\theta + \varepsilon} u(\rho^*(\mu), \tilde{\theta}) f(\tilde{\theta}) d\tilde{\theta} + \int_{\theta' - \varepsilon'}^{\theta' + \varepsilon'} u(\rho^*(\mu), \tilde{\theta}) f(\tilde{\theta}) d\tilde{\theta} = 0,$$

so that a posterior belief that pools the states $(\theta - \varepsilon, \theta + \varepsilon)$ and $(\theta' - \varepsilon', \theta' + \varepsilon')$ induces action $\rho^*(\mu)$. By continuity of u and V , this posterior belief is strictly preferred by the sender to full disclosure of these states.

Consider now the general case where $\theta, \theta' \in \text{supp}(\phi)$, but ϕ does not have atoms at θ and θ' nor does it have density in neighborhoods of θ and θ' . By Skorohod Representation Theorem, we can redefine the state so that it is uniformly distributed on $[0, 1]$. We can then use an argument similar to the one above to show that there exists a signal that is strictly preferred by the sender to full disclosure. \square

Proof of Lemma 3. Consider a pairwise signal τ . Suppose it is not single-dipped disclosure. Then there exist posterior beliefs $\mu, \mu' \in \text{supp}(\tau)$ that induce actions $\rho < \rho'$ and satisfy $\theta' \in (\theta, \theta'')$ for $\theta' \in \text{supp}(\mu')$ and $\{\theta, \theta''\} = \text{supp}(\mu)$.

Since posterior belief μ induces action $\rho < \rho' \leq 1$ and $u(\rho, \theta)$ satisfies upcrossing in θ , we have $u(\rho, \theta) \leq 0$. Consider first the trivial case where $u(\rho, \theta'') \leq 0$. A perturbation that splits μ into δ_θ and $\delta_{\theta''}$ (weakly) increases the sender's expected utility (since ρ cannot decrease) and changes a signal in the direction of single-dipped disclosure.

Consider now the remaining case where, by upcrossing of $u(\rho, \theta)$ in θ , we have

$$u(\rho, \theta) < 0 < u(\rho, \theta'').$$

Consider a perturbation that shifts weight $\hat{\alpha} \geq 0$ on θ from μ to μ' , shifts weight $\hat{\beta} \geq 0$ on θ' from μ' to μ , and shifts weight $\hat{\gamma} \geq 0$ on θ'' from μ to μ' .

This perturbation (i) increases the induced action ρ' if

$$\hat{\alpha}u(\rho', \theta) - \hat{\beta}u(\rho', \theta') + \hat{\gamma}u(\rho', \theta'') \geq 0,$$

(ii) increases the induced action ρ if

$$-\hat{\alpha}u(\rho, \theta) + \hat{\beta}u(\rho, \theta') - \hat{\gamma}u(\rho, \theta'') \geq 0,$$

and (iii) increases the sender's expected utility for fixed actions ρ and ρ' if

$$\hat{\alpha}(V(\rho', \theta) - V(\rho, \theta)) - \hat{\beta}(V(\rho', \theta') - V(\rho, \theta')) + \hat{\gamma}(V(\rho', \theta'') - V(\rho, \theta'')) \geq 0.$$

Observe that there exists a non-trivial perturbation, $(\hat{\alpha}, \hat{\beta}, \hat{\gamma}) \neq 0$, that satisfies conditions (i)–(iii) if and only if (5) holds. This perturbation increases the sender's expected utility and changes a signal τ in the direction of single-dipped disclosure.

To sum up, if (5) holds, then for each pairwise signal that is not single-dipped disclosure, we can change the signal in the direction of single-dipped disclosure and increase the sender's expected utility. Repeatedly applying such a perturbation yields single-dipped disclosure and completes the proof of the lemma. \square

Proof of Theorem 3. Only if. Suppose that, for $\theta < \theta' < \theta''$ and $\rho < \rho'$ such that $A < 0 < C$ and $c > 0$, there exists a vector $(\alpha, \beta, \gamma) \succeq 0$ that satisfies (5).

Since u satisfies downcrossing in ρ and $A < 0$, we have

$$\begin{aligned} a &< 0, \\ b \geq (>)0 &\implies B \geq (>)0. \end{aligned} \tag{22}$$

Since u satisfies aggregate downcrossing in ρ , Theorem A1 and Remark A1 yield

$$\begin{aligned} Ac &\geq Ca, \\ b \geq 0 &\implies Ab \geq Ba, \\ B \leq 0 &\implies Bc \geq Cb, \end{aligned} \tag{23}$$

where the first line holds because $A < 0 < C$ and $a < 0 < c$, the second line holds because $A < 0 \leq B$ and $a < 0 \leq b$ if $b \geq 0$, and the third line holds because $B \leq 0 < C$ and $b \leq 0 < c$ if $B \leq 0$.

By (22), there are three cases to consider: (1) $b \leq 0, B \geq 0$, (2) $b > 0, B > 0$, and (3) $b < 0, B < 0$. We denote $(x, y, z) = (|a|, |b|, |c|)$ and $(X, Y, Z) = (|A|, |B|, |C|)$, so that $(x, z) = (-a, c)$, $(X, Z) = (-A, C)$, and $Zx \geq Xz$ by (23).

(1) $b \leq 0, B \geq 0$. We have $y = -b, Y = B, Ab \geq 0 \geq Ba$, and $Bc \geq 0 \geq Cb$.

Consider first the subcase where $Zx > Xz$. Since there exists $(\alpha, \beta, \gamma) \succeq 0$ that satisfies (5), we have

$$\begin{aligned} \begin{cases} -x\alpha + y\beta + z\gamma \geq 0 \\ X\alpha + Y\beta - Z\gamma \geq 0 \\ \alpha - \beta + \gamma \geq 0 \end{cases} &\implies \begin{cases} \alpha \leq \frac{y\beta + z\gamma}{x} \\ \alpha \geq \frac{Z\gamma - Y\beta}{X} \\ \alpha \geq \beta - \gamma \end{cases} \implies \begin{cases} \frac{Z\gamma - Y\beta}{X} \leq \frac{y\beta + z\gamma}{x} \\ \beta - \gamma \leq \frac{y\beta + z\gamma}{x} \end{cases} \\ &\implies \begin{cases} (Zx - Xz)\gamma \leq (Xy + Yx)\beta \\ (x + z)\gamma \geq (x - y)\beta \end{cases} \implies \begin{cases} \gamma \leq \frac{Xy + Yx}{Zx - Xz}\beta \\ \gamma \geq \frac{x - y}{x + z}\beta \end{cases} \\ &\implies \frac{x - y}{x + z} \leq \frac{Xy + Yx}{Zx - Xz} \implies Ab - Ba + Bc - Cb \geq Ac - Ca. \end{aligned}$$

Thus, (6) holds in this subcase.

Consider now the subcase where $Zx = Xz$, so that $Ab \geq Ba$, $Bc \geq Cb$, and $Ac = Ca$. Thus, (6) trivially holds in this subcase.

(2) $b > 0, B > 0$. We have $y = b$, $Y = B$, and $Ab \geq Ba$, or equivalently $Yx \geq Xy$.

Consider first the subcase where $Yx > Xy$. Since there exists $(\alpha, \beta, \gamma) \succeq 0$ that satisfies (5), we have

$$\begin{aligned} \begin{cases} -x\alpha - y\beta + z\gamma \geq 0 \\ X\alpha + Y\beta - Z\gamma \geq 0 \\ \alpha - \beta + \gamma \geq 0 \end{cases} &\implies \begin{cases} \beta \leq \frac{z\gamma - x\alpha}{y} \\ \beta \geq \frac{Z\gamma - X\alpha}{Y} \\ \beta \leq \alpha + \gamma \end{cases} \implies \begin{cases} z\gamma - x\alpha \geq 0 \\ \frac{Z\gamma - X\alpha}{Y} \leq \frac{z\gamma - x\alpha}{y} \\ \frac{Z\gamma - X\alpha}{Y} \leq \alpha + \gamma \end{cases} \\ &\stackrel{(a)}{\implies} \begin{cases} x\alpha \leq z\gamma \\ (Yx - Xy)\alpha \leq (Yz - Zy)\gamma \\ (X + Y)\alpha \geq (Z - Y)\gamma \end{cases} \implies \begin{cases} \alpha \leq \frac{z\gamma}{x} \\ \alpha \leq \frac{Yz - Zy}{Yx - Xy}\gamma \\ \alpha \geq \frac{Z - Y}{X + Y}\gamma \end{cases} \\ &\implies \begin{cases} Yz - Zy \geq 0 \\ \frac{Z - Y}{X + Y} \leq \frac{Yz - Zy}{Yx - Xy} \end{cases} \implies \begin{cases} Bc \geq Cb \\ Ab - Ba + Bc - Cb \geq Ac - Ca. \end{cases} \end{aligned}$$

Thus, (6) holds in this subcase.

Consider now the subcase where $Yx = Xy$. Since there exists $(\alpha, \beta, \gamma) \succeq 0$ that satisfies (5), we have $Yz \geq Zy$, by the implication (a). Combining $Yx = Xy$, $Zx \geq Xz$, and $Yz \geq Zy$ implies that $Zx = Xz$, and $Yz = Zy$. Thus, (6) holds with equalities in this subcase.

(3) $b < 0, B < 0$. We have $y = -b$, $Y = -B$, and $Bc \geq Cb$, or equivalently $Zy \geq Yz$.

Consider first the subcase where $Zy > Yz$. Since there exists $(\alpha, \beta, \gamma) \succeq 0$ that satisfies (5), we have

$$\begin{aligned} \begin{cases} -x\alpha + y\beta + z\gamma \geq 0 \\ X\alpha - Y\beta - Z\gamma \geq 0 \\ \alpha - \beta + \gamma \geq 0 \end{cases} &\implies \begin{cases} \beta \geq \frac{x\alpha - z\gamma}{y} \\ \beta \leq \frac{X\alpha - Z\gamma}{Y} \\ \beta \leq \alpha + \gamma \end{cases} \implies \begin{cases} X\alpha - Z\gamma \geq 0 \\ \frac{x\alpha - z\gamma}{y} \leq \frac{X\alpha - Z\gamma}{Y} \\ \frac{x\alpha - z\gamma}{y} \leq \alpha + \gamma \end{cases} \end{aligned}$$

$$\begin{aligned}
& \stackrel{(b)}{\implies} \begin{cases} Z\gamma \leq X\alpha \\ (Zy - Yz)\gamma \leq (Xy - Yx)\alpha \\ (z + y)\gamma \geq (x - y)\alpha \end{cases} \implies \begin{cases} \gamma \leq \frac{X}{Z}\alpha \\ \gamma \leq \frac{Xy - Yx}{Zy - Yz}\alpha \\ \gamma \geq \frac{x - y}{z + y}\alpha \end{cases} \\
& \implies \begin{cases} Xy - Yx \geq 0 \\ \frac{x - y}{z + y} \leq \frac{Xy - Yx}{Zy - Yz} \end{cases} \implies \begin{cases} Ab \geq Ba \\ Ab - Ba + Bc - Cb \geq Ac - Ca. \end{cases}
\end{aligned}$$

Thus, (6) holds in this subcase.

Consider now the subcase where $Zy = Yz$. Since there exists $(\alpha, \beta, \gamma) \succeq 0$ that satisfies (5), we have $Xy \geq Yx$, by the implication (b). Combining $Zy = Yz$, $Zx \geq Xz$, and $Xy \geq Yx$ implies that $Zx = Xz$ and $Xy = Yx$. Thus, (6) holds with equalities in this subcase. \square

Proof of Corollary 3. We first notice that, in the canonical case, we only need to check the second line in condition (6). Indeed, consider a vector

$$(\alpha, \beta, \gamma) = (C - B, C - A, B - A).$$

In the canonical case, we have $A < B < C$ and thus $(\alpha, \beta, \gamma) \succeq 0$. The second and third inequalities in (5) hold with equality and the first inequality in (5) holds because it is equivalent to the second line in (6).

Let $\theta < \tilde{\theta} < \theta' < \tilde{\theta}' < \theta''$. If (7) holds, then

$$\begin{aligned}
& \int_{\theta'}^{\theta''} \int_{\theta}^{\theta'} \left(\frac{\partial u(\rho', \tilde{\theta}')}{\partial \theta} \frac{\partial u(\rho, \tilde{\theta})}{\partial \theta} - \frac{\partial u(\rho', \tilde{\theta})}{\partial \theta} \frac{\partial u(\rho, \tilde{\theta}')}{\partial \theta} \right) d\tilde{\theta} d\tilde{\theta}' \geq 0, \\
& \iff \int_{\theta'}^{\theta''} \int_{\theta}^{\theta'} \frac{\partial u(\rho', \tilde{\theta}')}{\partial \theta} \frac{\partial u(\rho, \tilde{\theta})}{\partial \theta} d\tilde{\theta} d\tilde{\theta}' \geq \int_{\theta'}^{\theta''} \int_{\theta}^{\theta'} \frac{\partial u(\rho', \tilde{\theta})}{\partial \theta} \frac{\partial u(\rho, \tilde{\theta}')}{\partial \theta} d\tilde{\theta} d\tilde{\theta}', \\
& \iff [u(\rho', \theta'') - u(\rho', \theta')][u(\rho, \theta') - u(\rho, \theta)] \geq [u(\rho, \theta'') - u(\rho, \theta')][u(\rho', \theta') - u(\rho', \theta)], \\
& \iff u(\rho, \theta)u(\rho', \theta') - u(\rho, \theta')u(\rho', \theta) + u(\rho, \theta')u(\rho', \theta'') - u(\rho, \theta'')u(\rho', \theta') \\
& \qquad \geq u(\rho, \theta)u(\rho', \theta'') - u(\rho, \theta'')u(\rho', \theta),
\end{aligned}$$

which is equivalent to the second line in (6) in the canonical case. So single-dipped disclosure is optimal. \square

Proof of Theorem 4. If. Consider a vector

$$(\alpha, \beta, \gamma) = \begin{cases} (c - b, c - a, b - a), & \text{if } a \leq b < c, \\ (C - B, C - A, B - A), & \text{otherwise.} \end{cases}$$

The first line in (9) implies that $(\alpha, \beta, \gamma) \succeq 0$.

Consider first the case $a \leq b < c$. The first and third inequalities in (8) hold with equality and the second inequality in (8) holds because it is equivalent to the third line in (9). That is, the perturbation holds fixed ρ' and the sender's expected utility (for fixed actions), while increasing ρ .

Consider now the case $A < B \leq C$. The second and third inequalities in (8) hold with equality and the first inequality in (8) holds because it is equivalent to the third line in (9). That is, the perturbation holds fixed ρ and the sender's expected utility (for fixed actions), while increasing ρ' .

Since (8) holds in both cases, single-peaked disclosure is optimal by Lemma 4.

Only if. The converse says that a perturbation that holds one of the two actions ρ and ρ' fixed and also fixes the sender's expected utility (for fixed ρ and ρ'), while increasing the other action, yields not only sufficient but also necessary conditions for single-peaked disclosure to be optimal in the case where actions ρ and ρ' are sufficiently close to each other, so that

$$u(\rho', \theta) < 0 \implies u(\rho, \theta) < 0.$$

It is possible provide the necessary and sufficient conditions for (8) to hold not only in the case $A < 0$ but also in the case $A \geq 0$, but since these conditions are quite complicated and are not used in our applications, we omit them.

Suppose that, for $\theta < \theta' < \theta''$ and $\rho < \rho'$ such that $a < 0 < c$ and $A < 0$, there exists a non-zero vector $(\alpha, \beta, \gamma) \succeq 0$ that satisfies (8).

Since u satisfies downcrossing in ρ and $c > 0$, we have

$$\begin{aligned} C &> 0, \\ b \geq (>)0 &\implies B \geq (>)0. \end{aligned} \tag{24}$$

Since u satisfies aggregate downcrossing in ρ , Theorem A1 and Remark A1 yield

$$\begin{aligned} Ac &\geq Ca, \\ b \geq 0 &\implies Ab \geq Ba, \\ B \leq 0 &\implies Bc \geq Cb, \end{aligned} \tag{25}$$

where the first line holds because $A < 0 < C$ and $a < 0 < c$, the second line holds because $A < 0 \leq B$ and $a < 0 \leq b$ if $b \geq 0$, and the third line holds because $B \leq 0 < C$ and $b \leq 0 < c$ if $B \leq 0$.

By (24), there are three cases to consider: (1) $b \leq 0, B \geq 0$, (2) $b > 0, B > 0$, and (3) $b < 0, B < 0$. We denote $(x, y, z) = (|a|, |b|, |c|)$ and $(X, Y, Z) = (|A|, |B|, |C|)$, so that $(x, z) = (-a, c)$, $(X, Z) = (-A, C)$, and $Zx \geq Xz$ by (25).

(1) $b \leq 0, B \geq 0$. We have $y = -b$, $Y = B$, and $A < 0 \leq B$.

Consider first the subcase where $Zx > Xz$. Since there exists $(\alpha, \beta, \gamma) \succeq 0$ that satisfies (8), we have

$$\begin{cases} -x\alpha + y\beta + z\gamma \leq 0 \\ X\alpha + Y\beta - Z\gamma \leq 0 \\ \alpha - \beta + \gamma \leq 0 \end{cases} \implies \begin{cases} \gamma \leq \frac{x\alpha - y\beta}{z} \\ \gamma \geq \frac{X\alpha + Y\beta}{Z} \\ \gamma \leq \beta - \alpha \end{cases} \implies \begin{cases} x\alpha - y\beta \geq 0 \\ \beta \geq \alpha \\ \frac{X\alpha + Y\beta}{Z} \leq \frac{x\alpha - y\beta}{z} \\ \frac{X\alpha + Y\beta}{Z} \leq \beta - \alpha \end{cases}$$

$$\begin{aligned}
&\stackrel{(a)}{\implies} \begin{cases} x\alpha \geq y\beta \\ \alpha \leq \beta \\ (Zx - Xz)\alpha \geq (Zy + Yz)\beta \\ (Z + X)\alpha \leq (Z - Y)\beta \end{cases} \implies \begin{cases} \alpha \geq \frac{y}{x}\beta \\ \alpha \leq \beta \\ \alpha \geq \frac{Zy + Yz}{Zx - Xz}\beta \\ \alpha \leq \frac{Z - Y}{Z + X}\beta \end{cases} \\
&\implies \begin{cases} Z - Y \geq 0 \\ \frac{Zy + Yz}{Zx - Xz} \leq \frac{Z - Y}{Z + X} \end{cases} \implies \begin{cases} B \leq C \\ Ab - Ba + Bc - Cb \leq Ac - Ca. \end{cases}
\end{aligned}$$

Thus, (9) holds in this subcase.

Consider now the subcase where $Zx = Xz$. Since there exists $(\alpha, \beta, \gamma) \succeq 0$ that satisfies (8), we have $Zy + Yz = 0$, by the implication (a). Thus, $b = B = 0$ and (9) trivially holds in this subcase.

(2) $b > 0, B > 0$. We have $y = b, Y = B$, and $A < 0 < B$.

Consider first the subcase where $Zx > Xz$. Since there exists $(\alpha, \beta, \gamma) \succeq 0$ that satisfies (8), we have

$$\begin{aligned}
&\begin{cases} -x\alpha - y\beta + z\gamma \leq 0 \\ X\alpha + Y\beta - Z\gamma \leq 0 \\ \alpha - \beta + \gamma \leq 0 \end{cases} \implies \begin{cases} \gamma \leq \frac{x\alpha + y\beta}{z} \\ \gamma \geq \frac{X\alpha + Y\beta}{Z} \\ \gamma \leq \beta - \alpha \end{cases} \implies \begin{cases} \beta \geq \alpha \\ \frac{X\alpha + Y\beta}{Z} \leq \frac{x\alpha + y\beta}{z} \\ \frac{X\alpha + Y\beta}{Z} \leq \beta - \alpha \end{cases} \\
&\stackrel{(b)}{\implies} \begin{cases} \alpha \leq \beta \\ (Zx - Xz)\alpha \geq (Yz - Zy)\beta \\ (Z + X)\alpha \leq (Z - Y)\beta \end{cases} \implies \begin{cases} \alpha \leq \beta \\ \alpha \geq \frac{Yz - Zy}{Zx - Xz}\beta \\ \alpha \leq \frac{Z - Y}{Z + X}\beta \end{cases} \\
&\implies \begin{cases} Z - Y \geq 0 \\ \frac{Yz - Zy}{Zx - Xz} \leq \frac{Z - Y}{Z + X}, \end{cases} \implies \begin{cases} B \leq C \\ Ab - Ba + Bc - Cb \leq Ac - Ca. \end{cases}
\end{aligned}$$

Thus, (9) holds in this subcase.

Consider now the subcase where $Zx = Xz$, or equivalently $Ac = Ca$. Since there exists $(\alpha, \beta, \gamma) \succeq 0$ that satisfies (5), we have $Z \geq Y$ and $Yz \leq Zy$, or equivalently $B \leq C$ and $Bc \leq Cb$, by the implication (b). So

$$Ab - Ba + Bc - Cb = \frac{c - a}{c}(Bc - Cb) \leq 0 = Ac - Ca,$$

where the first equality holds by $Ac = Ca$, the inequality holds by $a < 0 < c$ and $Bc \leq Cb$, and the last equality holds by $Ac = Ca$. Thus, (9) holds in this subcase.

(3) $b < 0, B < 0$. We have $y = -b, Y = -B, b < 0 < c$, and $Bc \geq Cb$, or equivalently $Zy \geq Yz$.

Consider first the subcase where $Zy > Yz$. Since there exists $(\alpha, \beta, \gamma) \succeq 0$ that satisfies (8), we have

$$\begin{aligned} \begin{cases} -x\alpha + y\beta + z\gamma \leq 0 \\ X\alpha - Y\beta - Z\gamma \leq 0 \\ \alpha - \beta + \gamma \leq 0 \end{cases} &\implies \begin{cases} \beta \leq \frac{x\alpha - z\gamma}{y} \\ \beta \geq \frac{X\alpha - Z\gamma}{Y} \\ \beta \geq \alpha + \gamma \end{cases} \implies \begin{cases} x\alpha - z\gamma \geq 0 \\ \frac{X\alpha - Z\gamma}{Y} \leq \frac{x\alpha - z\gamma}{y} \\ \alpha + \gamma \leq \frac{x\alpha - z\gamma}{y} \end{cases} \\ &\stackrel{(c)}{\implies} \begin{cases} z\gamma \leq x\alpha \\ (Zy - Yz)\gamma \geq (Xy - Yx)\alpha \\ (z + y)\gamma \leq (x - y)\alpha \end{cases} \implies \begin{cases} \gamma \leq \frac{x}{z}\alpha \\ \gamma \geq \frac{Xy - Yx}{Zy - Yz}\alpha \\ \gamma \leq \frac{x - y}{z + y}\alpha \end{cases} \\ &\implies \begin{cases} x - y \geq 0 \\ \frac{Xy - Yx}{Zy - Yz} \leq \frac{x - y}{y + z} \end{cases} \implies \begin{cases} a \leq b \\ Ab - Ba + Bc - Cb \leq Ac - Ca. \end{cases} \end{aligned}$$

Thus, (9) holds in this subcase.

Consider now the subcase where $Zy = Yz$, or equivalently $Bc = Cb$. Since there exists $(\alpha, \beta, \gamma) \succeq 0$ that satisfies (8), we have $x \geq y$ and $Xy \leq Yx$, or equivalently $a \leq b$ and $Ab \leq Ba$, by the implication (c). So

$$Ab - Ba + Bc - Cb \leq 0 \leq Ac - Ca,$$

where the first inequality holds by $Ab \leq Ba$ and $Bc = Cb$, and the last inequality holds by $Ac \geq Ca$. Thus, (9) holds in this subcase. \square

Proof of Proposition 4. Consider a vector

$$(\hat{\alpha}, \hat{\beta}, \hat{\gamma}) = \begin{cases} (0, \theta'' - \rho', \theta' - \rho'), & \text{if } \rho' \leq \theta', \\ (\rho' - \theta', \rho' - \theta, 0), & \text{if } \rho' > \theta'. \end{cases}$$

The inequality $u(\rho, \theta) < 0 < u(\rho, \theta'')$ is equivalent to $\theta < \rho < \theta''$. Moreover, since $\theta < \theta' < \theta''$, we have $(\hat{\alpha}, \hat{\beta}, \hat{\gamma}) \succeq 0$.

Consider first the case $\rho' \leq \theta'$. The first inequality in (11) holds with equality. The second inequality in (11) holds because

$$(\theta' - \rho)(\theta'' - \rho') - (\theta'' - \rho)(\theta' - \rho') = (\rho' - \rho)(\theta'' - \theta') > 0.$$

The third inequality in (11) holds with equality because $\rho < \rho' \leq \theta' < \theta''$ and thus $V(\rho', \theta') = V(\rho, \theta') = V(\rho', \theta'') = V(\rho, \theta'') = 0$.

Consider now the case $\rho' > \theta'$. The first inequality in (11) holds with equality. The second inequality in (11) holds because

$$-(\theta - \rho)(\rho' - \theta') + (\theta' - \rho)(\rho' - \theta) = (\rho' - \rho)(\theta' - \theta) > 0.$$

Thus, it suffices to show that the third inequality in (11) holds,

$$(V(\rho', \theta) - V(\rho, \theta))(\rho' - \theta') - (V(\rho', \theta') - V(\rho, \theta'))(\rho' - \theta) \geq 0,$$

or equivalently,

$$\frac{V(\rho', \theta) - V(\rho, \theta)}{\rho' - \theta} \geq \frac{V(\rho', \theta') - V(\rho, \theta')}{\rho' - \theta'}.$$

That is, it suffices to show that

$$H(\theta') = \frac{V(\rho', \theta') - V(\rho, \theta')}{\rho' - \theta'}$$

is decreasing in θ' for $\theta' < \rho'$.

There are two subcases to consider. First, if $\rho \leq \theta' < \rho'$, then $V(\rho, \theta') = 0$, and

$$H(\theta') = (\rho' - \theta')^\zeta,$$

which is decreasing in θ' .

Second, if $\theta' < \rho$, then $V(\rho, \theta') = (\rho - \theta')^{\zeta+1}$, and the derivative of H is

$$\frac{dH(\theta')}{d\theta'} = \frac{-\zeta(\rho' - \theta')^{\zeta+1} + (\zeta + 1)(\rho - \theta')^\zeta(\rho' - \theta') - (\rho - \theta')^{\zeta+1}}{(\rho' - \theta')^2},$$

which is negative if

$$h(t) = -\zeta t^{\zeta+1} + (\zeta + 1)t - 1 < 0 \text{ for } t > 1,$$

which holds because $h(1) = 0$ and

$$\frac{dh(t)}{dt} = (\zeta + 1)(1 - \zeta t^\zeta) < 0 \text{ for } t > 1 \text{ and } \zeta \geq 1. \quad \square$$

A.3. Threshold-Gain Preferences. Here we consider a special case we call *threshold gain* preferences. Threshold gain preferences are somewhat similar to the overvaluation preferences considered in Section 4.4. We are interested in these preferences primarily because they allow us to nest the main result of [Goldstein and Leitner \(2018\)](#) in our model.

Assume that $u(\rho, \theta) = \theta - \rho$ and

$$V(\rho, \theta) = \begin{cases} w(\theta)\mathbb{1}_{\rho \geq \theta_0}, & \text{if } \rho \geq \sigma(\theta), \\ -\infty, & \text{otherwise,} \end{cases}$$

where $\theta_0 \in (0, 1)$ is a constant, w is a continuous, strictly positive function, and σ is a continuous, increasing function that satisfies $\sigma(\theta) = \theta$ for $\theta \in [0, \theta_0]$ and $\sigma(\theta) < \theta$ for $\theta \in (\theta_0, 1]$. Notice that full disclosure guarantees a positive expected utility.

The model of [Goldstein and Leitner \(2018\)](#) can be simplified to this setting. The sender is a bank regulator and the receiver is a perfectly competitive market. The bank has an asset which yields a random cash flow. The timing is as follows: The bank privately observes the expected cash flow θ . The regulator designs a stress test to reveal information about θ . After observing the test results, the market offers a fair price ρ for the asset. The bank decides whether to keep the asset and receive the random cash flow or sell it at the price ρ .

The indicator function $\mathbb{1}_{\rho \geq \theta_0}$ in the regulator's utility captures the idea that there is a gain when the bank's cash holdings exceed a threshold θ_0 . Formally, the market price should reflect not only the information revealed by the regulator but also the anticipated decision of the bank as to whether to sell the asset. We implicitly incorporate the bank's decision by restricting attention, without loss of generality, to tests that always induce the bank to sell the asset. This is done by specifying that the regulator gets infinitely low utility if the chosen test ever generates a price ρ for the asset θ that falls below its reservation price $\sigma(\theta)$ at which the bank is willing to sell the asset.

When $\theta > \theta_0$, the bank strictly prefers to sell the asset at price θ than to hold it and risk cash holdings falling below θ_0 . Hence, $\theta_0 \leq \sigma(\theta) < \theta$ for $\theta > \theta_0$. When $\theta < \theta_0$, the bank strictly prefers to hold the asset and gamble that cash holdings exceed θ_0 , rather than selling at the price θ . Hence, the asset is never sold at the price below θ_0 , and it is irrelevant what information about θ is revealed in this case. So, without loss of generality, we can redefine the reservation price to be $\sigma(\theta) = \theta$ for $\theta < \theta_0$, so that the regulator is forced to reveal the asset's type θ whenever the induced price is below θ_0 to make the bank willing to sell the asset.

With these preferences, $V(\rho, \theta)$ is not smooth and is not strictly increasing in ρ and thus it violates Assumptions 1 and 2. Nevertheless, Lemma 3 still applies in the following form (with essentially the same proof):

Lemma A3. *Single-dipped disclosure is optimal if, for all $\theta < \theta' < \theta''$ and $\rho < \rho'$ such that $u(\rho, \theta) < 0 < u(\rho, \theta'')$ and $V(\rho, \theta'') \neq -\infty$, there exists a non-zero vector $(\hat{\alpha}, \hat{\beta}, \hat{\gamma}) \geq 0$ such that*

$$\begin{pmatrix} u(\rho', \theta) & -u(\rho', \theta') & u(\rho', \theta'') \\ -u(\rho, \theta) & u(\rho, \theta') & -u(\rho, \theta'') \\ V(\rho', \theta) - V(\rho, \theta) & -(V(\rho', \theta') - V(\rho, \theta')) & V(\rho', \theta'') - V(\rho, \theta'') \end{pmatrix} \begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \\ \hat{\gamma} \end{pmatrix} \geq \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \quad (26)$$

Lemma A3 yields the following result, which reproduces Goldstein and Leitner (2018, Proposition 3).

Proposition A1. *In the threshold-gain preferences case, unitary single-dipped disclosure is optimal.*

Proof. Notice that $V(\rho, \theta'') \neq -\infty$ is equivalent to $\rho \geq \sigma(\theta'')$ and $u(\rho, \theta) < 0 < u(\rho, \theta'')$ is equivalent to $\theta < \rho < \theta''$. Taking into account that $\sigma(\theta) = \theta$ for $\theta \leq \theta_0$ gives $\rho > \theta_0$. Thus the last row in (26) is zero, and the third inequality in (26) always holds with equality.

Consider a vector

$$(\hat{\alpha}, \hat{\beta}, \hat{\gamma}) = \begin{cases} (0, \theta'' - \rho', \theta' - \rho'), & \text{if } \rho' \leq \theta', \\ (\rho' - \theta', \rho' - \theta, 0), & \text{if } \rho' > \theta'. \end{cases}$$

The proof of Proposition 4 shows that the first inequality in (26) holds with equality and the second inequality in (26) holds with strict inequality. Thus, single-dipped disclosure is optimal by Lemma A3.

It is now easy to see that *unitary* single-dipped disclosure is optimal. Otherwise, there are paired states on the same side of θ_0 , and the sender's expected utility (weakly) increases if these states are separated. \square

REFERENCES

- AUMANN, R. J., AND M. MASCHLER (1995): *Repeated Games with Incomplete Information*. MIT press.
- BAGNOLI, M., AND T. BERGSTROM (2005): "Log-Concave Probability and Its Applications," *Economic theory*, 26(2), 445–469.
- BECKER, G. S. (1973): "A Theory of Marriage: Part I," *Journal of Political Economy*, 81(4), 813–846.
- CANDOGAN, O. (2019): "Optimality of Double Intervals in Persuasion: A Convex Programming Framework," University of Chicago.
- CHOI, M., AND L. SMITH (2017): "Ordinal Aggregation Results via Karlin's Variation Diminishing Property," *Journal of Economic Theory*, 168, 1–11.
- DWORCZAK, P., AND A. KOLOTILIN (2019): "The Persuasion Duality," Northwestern University.
- DWORCZAK, P., AND G. MARTINI (2019): "The Simple Economics of Optimal Persuasion," *Journal of Political Economy*, 127(5), 1993–2048.
- FRIEDMAN, J. N., AND R. T. HOLDEN (2008): "Optimal Gerrymandering: Sometimes Pack, but Never Crack," *American Economic Review*, 98(1), 113–44.
- GENTZKOW, M., AND E. KAMENICA (2016): "A Rothschild-Stiglitz Approach to Bayesian Persuasion," *American Economic Review, Papers & Proceedings*, 106, 597–601.
- GOLDSTEIN, I., AND Y. LEITNER (2018): "Stress Tests and Information Disclosure," *Journal of Economic Theory*, 177, 34–69.
- GUO, Y., AND E. SHMAYA (2019): "The Interval Structure of Optimal Disclosure," *Econometrica*, 87(2), 653–675.
- KAMENICA, E., AND M. GENTZKOW (2011): "Bayesian Persuasion," *American Economic Review*, 101, 2590–2615.
- KARLIN, S., AND H. RUBIN (1956): "The Theory of Decision Procedures for Distributions with Monotone Likelihood Ratio," *Annals of Mathematical Statistics*, 27(2), 272–299.
- KOLOTILIN, A. (2018): "Optimal Information Disclosure: A Linear Programming Approach," *Theoretical Economics*, 13, 607–636.
- KOLOTILIN, A., T. MYLOVANOV, A. ZAPECHELNYUK, AND M. LI (2017): "Persuasion of a Privately Informed Receiver," *Econometrica*, 85, 1949–1964.
- KREMER, M., AND E. MASKIN (1996): "Wage Inequality and Segregation by Skill," Harvard University.

- LEGROS, P., AND A. F. NEWMAN (2002): “Monotone Matching in Perfect and Imperfect Worlds,” *Review of Economic Studies*, 69(4), 925–942.
- NIKANDROVA, A., AND R. PANCS (2017): “Conjugate Information Disclosure in an Auction with Learning,” *Journal of Economic Theory*, 171, 174–212.
- QUAH, J., AND B. STRULOVICI (2012): “Aggregating the Single Crossing Property,” *Econometrica*, 80, 2333–2348.
- RAYO, L., AND I. SEGAL (2010): “Optimal Information Disclosure,” *Journal of Political Economy*, 118, 949–987.
- TERSTIEGE, S., AND C. WASSER (2019): “Buyer-Optimal Extensionproof Information,” University of Bonn.