# Online Appendices

## D. Examples of Negative Assortative Disclosure

**Example 1** (Solving the Differential Equations)**.** Consider the linear receiver case with $Y = X = [1/e, e]$, $f(x) = 1/(2x)$, and $V(y, x) = y/x$. We claim that the unique optimal outcome matches each state $x_1 \in [1/e, 1]$ with state $x_2 = 1/x_1$ with equal weights, so that the induced action is $y = x_1/2 + 1/(2x_1)$. Thus, $\chi_1(y) = y - \sqrt{y^2 - 1}$, and $\chi_2(y) = y + \sqrt{y^2 - 1}$ for all $y \in [1, e/2 + 1/(2e)]$.

Indeed, by Theorem 4, the optimal outcome is strictly single-dipped, since $w(x) = 1/x$ is strictly convex. By Corollary 3, (12) holds, since $w' < 0$. Hence, by Theorem 6, the optimal outcome is single-dipped negative assortative and satisfies (9)–(11). Now, for $x_2 = 1/x_1$ and $y = x_1/2 + 1/(2x_1)$, (9) holds because

$$u(y, x_2) = \left(\frac{1}{2x_1} - \frac{x_1}{2}\right) = -\left(\frac{x_1}{2} - \frac{1}{2x_1}\right) = -u(y, x_1),$$
$$f(x_2)\frac{\mathrm{d}x_2}{\mathrm{d}y} = \frac{1}{2/x_1}\left(-\frac{1}{x_1^2}\frac{\mathrm{d}x_1}{\mathrm{d}y}\right) = -\frac{1}{2x_1}\frac{\mathrm{d}x_1}{\mathrm{d}y} = -f(x_1)\frac{\mathrm{d}x_1}{\mathrm{d}y},$$

(10) holds because

$$\frac{\mathrm{d}}{\mathrm{d}y}\left(w(x_1)\frac{1}{2} + w(x_2)\frac{1}{2}\right) = \frac{\mathrm{d}}{\mathrm{d}y}\left(\frac{1}{2x_1} + \frac{x_1}{2}\right) = \frac{\mathrm{d}}{\mathrm{d}y}y = 1,$$
$$\frac{w(x_2) - w(x_1)}{x_2 - x_1} = \frac{x_1 - 1/x_1}{1/x_1 - x_1} = -1,$$

and (11) holds because $1/(1/e) = e$ and $1/1 = 1$.[56]

**Example 2** (Quantile Persuasion)**.** Consider the quantile sub-subcase of the state-independent sender case, where $u(y, x) = 1\{x \geq y\} - \kappa$ with $\kappa \in (0, 1)$. Let $\phi$ have a density on $[0, 1]$. Assuming that the receiver breaks ties in favor of the sender, we have, for $x_1 < x_2$,

$$\gamma(\rho\delta_{x_1} + (1-\rho)\delta_{x_2}) = \begin{cases} x_2, & \rho \leq 1 - \kappa, \\ x_1, & \rho > 1 - \kappa. \end{cases}$$

Note that (12) always holds for $\rho \in (0, 1 - \kappa)$. We claim that there exists an optimal single-dipped negative assortative signal where the induced distribution over actions $\alpha$ satisfies $\alpha([y, 1]) = \phi([y, 1])/\kappa$, and the posterior inducing any action $y \in [\underline{y}, 1]$ is

[56] We can also solve this example by directly applying Theorem 1, because, for $q(y) = y$, the function $V(y, x) + q(y)u(y, x) = y/x + y(x - y)$ is maximized at $y = x/2 + 1/(2x)$ for all $x \in [1/e, e]$.

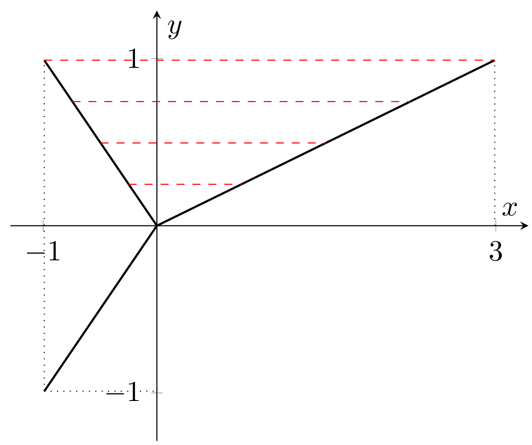


FIGURE 5. The Optimal Signal in Example 3

*Notes:* The three solid line segments depict the graphs of $\chi_1$ and $\chi_2$. The dashed line segments indicate pairs of states $\chi_1(y)$ and $\chi_2(y)$ that are optimal to pool to induce action $y \in (0,1]$. If the prior density satisfies $f(-y) > 3f(3y)$ for all $y \in (0,1]$, then, for each state $x < 0$, the optimal signal randomizes between disclosing $x$ (and inducing action $x$) and pooling $x$ with state $-3x$ (and inducing action $-x$).

$(1-\kappa)\delta_{\chi_1(y)} + \kappa\delta_{\chi_2(y)}$, where $\chi_2(y) = y$, $\chi_1(y)$ solves $\kappa\phi([0,\chi_1(y)]) = (1-\kappa)\phi([y,1])$, and $\underline{y}$ solves $\kappa\phi([0,\underline{y}]) = (1-\kappa)\phi([\underline{y},1])$.[57] A notable feature of this signal is that, with the informed receiver interpretation, it would remain optimal even if the sender knew the receiver's type and could condition disclosure on it.

**Example 3** (A Stochastic Optimal Signal)**.** In the following example, for some priors negative assortative disclosure is optimal; and for other priors, the unique optimal signal randomizes conditional on the state, even though the prior is atomless.

Consider the translation-invariant subcase of the state-independent sender case. Let $Y = X = [-1,3]$, let $\phi$ have a density $f$ with $f(-y) \geq 3f(3y)$ for all $y \in (0,1]$, let $u(y,x) = T(x-y)$ with $T(0) = 0$ and strictly log-concave $T'$, and let $V(y,x) = T(2y)$. With the informed receiver interpretation, this captures a case where, for example, $\kappa = 1/2$, the distribution of $\varepsilon$ is $N(0,\sigma^2)$, and the distribution of $t$ is $N(0,(\sigma/2)^2)$.[58]

[57] See Appendix F.18 for the proof.

[58] By symmetry and strict log-concavity of $T'$, $V_{yy}(y)/V_y(y) = 2T''(2y)/T'(2y) > (<) T''(0)/T'(0) = 0$ for $0 > (<) y$, showing that (13) fails for $y < 0$, and thus Theorem 6 does not apply.

By Theorem 4, the optimal outcome is strictly single-dipped. Furthermore, we claim that

$$\chi_1(y) = \begin{cases} y, & y \in [-1,0], \\ -y, & y \in (0,1], \end{cases} \quad \text{and} \quad \chi_2(y) = \begin{cases} y, & y \in [-1,0], \\ 3y, & y \in (0,1], \end{cases}$$

so that the posterior inducing any action $y \in [-1,1]$ is $\delta_{\chi_1(y)}/2 + \delta_{\chi_2(y)}/2$, and the distribution over actions $\alpha$ has density $a$ given by

$$a(y) = \begin{cases} 6f(3y), & y \in (0,1], \\ f(-y) - 3f(3y), & y \in [-1,0). \end{cases}$$

The unique optimal signal is single-dipped negative assortative iff $f(-y) = 3f(3y)$ for all $y \in (0,1]$. In contrast, if $f(-y) > 3f(3y)$ for all $y \in (0,1]$, then each state $x \in [-1,0)$ is mixed between inducing actions $y = x$ and $y = -x$.[59] See Figure 5.

## E. Specific Persuasion Models

This appendix shows how our analysis covers some well-known prior persuasion models, where single-dipped or single-peaked disclosure is optimal.[60]

E.1. **Contests.** Zhang and Zhou (2016) study information disclosure in contests. In their model, two contestants, $A$ and $B$, compete for a prize by exerting efforts $z_A$ and $z_B$. The probability that contestant $i = A, B$ wins is $z_i/(z_A + z_B)$. Everyone knows contestant $A$'s value $v_A = 1$. Contestant $B$'s value $v_B$ is known to contestant $B$ and the designer. The sender designs a signal about $v_B$ to maximize expected total effort.

It is convenient to parameterize $x = 1/\sqrt{v_B}$ and $y = \sqrt{z_A}$. With this parameterization, Zhang and Zhou's Proposition 1 shows that, given a posterior $\mu$, contestant $A$ exerts effort $z_A^\star = \gamma(\mu)^2$ determined by $\mathbb{E}_\mu\left[x - (1 + x^2)\,\gamma(\mu)\right] = 0$, and contestant $B$ (who knows $x$) exerts effort $z_B^\star(x) = \gamma(\mu)/x - \gamma(\mu)^2$, so the sender's expected utility is $z_A^\star + \mathbb{E}_\mu\left[z_B^\star(x)\right] = \mathbb{E}_\mu\left[\gamma(\mu)/x\right]$. We thus recover our model with $V(y,x) = y/x$ and $u(y,x) = x - (1+x^2)y$.

Zhang and Zhou give results on optimality of pairwise disclosure, full disclosure, and no disclosure. Our approach easily yields the following result, which additionally gives

[59] See Appendix F.19 for the proof.

[60] The applications in this appendix also illustrate some technical points. Appendix E.1 illustrates how directly applying Theorem 3 can yield weaker sufficient conditions for the optimality of single-dipped disclosure than those in Theorem 4. Appendices E.2 and E.3 illustrate how our analysis extends when some of our assumptions are violated: in Appendix E.2, Assumption 3 fails, so the receiver's optimal action may be at the boundary and thus violate the first-order condition; in Appendix E.3, Assumption 4 fails, as the sender only weakly prefers higher actions.

conditions for optimality of single-dipped/-peaked disclosure and negative assortative disclosure (which were not considered by Zhang and Zhou).

**Proposition 2.** *In Zhang and Zhou's contest model where the prior $\phi$ has a positive density on $X = [\underline{x}, \overline{x}]$, where $0 < \underline{x} < \overline{x}$, if $\underline{x} \geq 1$ then the unique optimal signal is full disclosure; and if $\overline{x} \leq 1/\sqrt{3}$ ($1/\sqrt{3} \leq \underline{x} < \underline{x} \leq 1$), then the unique optimal signal is single-dipped (-peaked) negative assortative disclosure.*

The proof of single-dippedness/-peakedness uses Theorem 3 with a perturbation that fixes both actions. In contrast, directly applying Theorem 4 would yield only the weaker result that single-peaked negative assortative disclosure is optimal if $1/\sqrt{2} \leq \underline{x} < \overline{x} < 1$.[61]

E.2. **Affiliated Information.** Guo and Shmaya (2019) consider a persuasion model with a privately informed receiver, where it is commonly known that the receiver wishes to accept a proposal iff $x$ exceeds a threshold $x_0$, and the receiver's type $t$ is his private signal of $x$. Letting $G(t|x)$ denote the distribution of $t$ conditional on $x$, with corresponding density $g(t|x)$, this setup maps to our model with $V(y, x) = G(y|x)$, $u(y, x) = (x - x_0)g(y|x)$, and $g(t|x)$ strictly log-submodular in $(t, x)$.[62,63] These preferences satisfy Assumptions 1 and 2 (see Lemma 3), but not Assumption 3, as $u(y, x) > 0$ for all $y$ when $x > x_0$. Nonetheless, assuming that the receiver breaks ties in the sender's favor, we have $\gamma(\mu) = \max\{y : \mathbb{E}_\mu[u(y, x)] \geq 0\}$.

Let us take for granted that Theorem 3 holds even though Assumption 3 is violated (e.g., this is clearly true with a discrete prior). Applying Theorem 3 with a perturbation that fixes one action while increasing the other action and the sender's expected utility (for fixed actions), we obtain the following result, which reproduces Guo and Shmaya's main qualitative insight.[64]

**Proposition 3.** *In Guo and Shmaya's model of persuading a privately informed receiver, every optimal signal is single-peaked.*

[61] To see this, suppose $\overline{x} < 1$. Then $u_x(y, x) = 1 - 2xy > 0$ for $y \leq \overline{x}/(1 + \overline{x}^2) = \max Y$. Moreover, $u_{yx}(y, x)/u_x(y, x) = -2x/(1 - 2xy)$ is always decreasing in $x$, while $V_{yx}(y_2, x)/u_x(y_1, x) = -1/(x^2 - 2x^3 y_1)$ is decreasing in $x$ iff $3\underline{x} \min Y = 3\underline{x}^2/(1 + \underline{x}^2) \geq 1$, or equivalently $\underline{x} \geq 1/\sqrt{2}$.

[62] The ordering convention here is that high $t$ is bad news about $x$. This ordering is opposite to Guo and Shmaya's, but follows our convention that the receiver accepts for types below a cutoff.

[63] Inostroza and Pavan (2025) study robust stress test design in a setting with multiple receivers with coordination motives. As they note, the single-receiver version of their model is a special case of Guo and Shmaya (2019).

[64] When the prior has positive density on $[0, 1]$, Guo and Shmaya's Theorem 3.1 additionally implies that the optimal signal is single-peaked negative assortative.

E.3. **Stress Tests.** Goldstein and Leitner (2018) consider a model of optimal stress tests. The sender is a bank regulator and the receiver is a perfectly competitive market. The bank has an asset that yields a random cash flow. The asset's quality is $x$, which is observed by the bank and the regulator but not the market, and is normalized to equal the asset's expected cash flow.[65] The regulator designs a test to reveal information about $x$. After observing the test result, the market offers a competitive price $y$ for the asset. Finally, the bank decides whether to keep the asset and receive the random cash flow, or sell it at price $y$. Letting $z$ denote the bank's final cash holding (equal to either the random cash flow or $y$), the bank's payoff equals $z + 1\{z \geq x_0\}$, where $x_0$ is a constant. An interpretation is that the bank faces a run if its cash holding falls below $x_0$. The regulator designs the test to maximize expected social welfare, or equivalently to minimize the probability of a run.

Goldstein and Leitner show that a bank with a type-$x$ asset is willing to sell at a price $y$ iff $y$ exceeds a reservation price $\tilde{\sigma}(x)$ that satisfies $\tilde{\sigma}(x) > x$ if $x < x_0$, $\tilde{\sigma}(x) < x$ if $x > x_0$, and $\tilde{\sigma}'(x) \geq 0$. Intuitively, if $x < x_0$ then the bank demands a premium to forego the chance that a lucky cash flow shock pushes its holdings above $x_0$, while if $x > x_0$ then the bank desires insurance against bad cash flow shocks that push its holdings below $x_0$. However, the value of the regulator's problem is unaffected if the reservation price is re-defined as $\sigma(x) = x$ if $x \leq x_0$ and $\sigma(x) = \tilde{\sigma}(x)$ if $x > x_0$, because it is suboptimal for the regulator to induce a bank to sell at a price below $x_0$. It is more convenient to work with the normalized reservation price $\sigma(x)$.

It is also convenient to restrict attention to tests that, for each $x$, either induce the bank to sell or fully disclose the bank's value: this is without loss because pooling two asset types that do not sell is weakly worse than disclosing these types. For such a test, the price induced by any posterior $\mu$ is $\gamma(\mu) = \mathbb{E}_\mu[x]$, so we are in the linear receiver case. We can capture the requirement that the bank always sells if $y \neq x$ by setting $V(y, x) = -\infty$ if $y < \sigma(x)$. Finally, letting $w(x) > 0$ equal the social gain when a bank sells a type-$x$ asset at a price above $x_0$ (which equals the probability that a type-$x$ asset yields a cash flow below $x_0$), we obtain the linear receiver case of our model with

$$V(y, x) = \begin{cases} w(x)1\{y \geq x_0\}, & \text{if } y \geq \sigma(x), \\ -\infty, & \text{otherwise.} \end{cases}$$

[65] This is the model in Section 5 of their paper, where the bank observes $x$.

Note that $V$ violates Assumptions 1 and 4, as it is discontinuous and only weakly increasing in $y$. Nonetheless, if we assume a discrete prior (as do Goldstein and Leitner), we recover their main qualitative insight.

**Proposition 4.** *In Goldstein and Leitner's stress test model with a discrete prior, there exists an optimal single-dipped signal.*

To prove the proposition, we use a perturbation that fixes both actions. Since $V$ is only weakly increasing, this perturbation now only weakly increases the sender's expected utility. Nonetheless, when $X$ is finite, repeatedly applying such perturbations eventually yields a single-dipped signal. We also note that, as Goldstein and Leitner show, if $\mathbb{E}_\phi[x] < x_0$—so that no disclosure does not attain the sender's first-best outcome—then every optimal signal is single-dipped.[66]

## F. Additional Proofs

F.1. **Proof of Lemma 3.** 1 $\implies$ 2. It is easy to see that Assumption 2 for $\mu = \delta_x$ such that $u(y, x) = 0$ yields (14). Similarly, Assumption 2 for $\mu = \rho\delta_x + (1-\rho)\delta_x$ such that $u(y, x) < 0 < u(y, x')$ and $\rho u(y, x) + (1-\rho)u(y, x') = 0$ yields (15).

2 $\implies$ 1. By Lemma 11, for any $y \in Y$ and $\mu \in \Delta(X)$ such that $\int u(y, x)\mathrm{d}\mu(x) = 0$, there exists $\sigma_\mu \in \Delta(\Delta(X))$ such that $\int \eta \mathrm{d}\sigma_\mu(\eta) = \mu$, and for each $\eta \in \mathrm{supp}(\sigma_\mu)$ there exist $x, x' \in X$ and $\rho \in [0, 1]$ such that $\eta = \rho\delta_x + (1-\rho)\delta_{x'}$ and

$$\rho u(y, x) + (1-\rho)u(y, x') = 0. \tag{19}$$

It suffices to show that

$$\rho u_y(y, x) + (1-\rho)u_y(y, x') < 0. \tag{20}$$

There are two cases to consider. First, if $\rho u(y, x) = 0$, then (20) follows from (14) and (19). Second, if $\rho u(y, x) \neq 0$, then (20) follows from (15) and (19).

3 $\implies$ 1. Notice that

$$\int u(y, x)\mathrm{d}\mu(x) = 0 \iff \int \tilde{u}(y, x)\mathrm{d}\mu(x) = 0.$$

[66] A related model by Garcia and Tsur (2021) studies optimal information disclosure to facilitate trade in an insurance market with adverse selection. Their model can be mapped to the linear receiver case with $V(y, x) = \nu(y)$ if $y \geq \sigma(x)$ and $V(y, x) = -\infty$ otherwise, where $\nu(y)$ is a strictly increasing, strictly concave function, and $\sigma$ is a continuous, strictly increasing function that satisfies $\sigma(x) < x$. Considering a similar perturbation as in Goldstein and Leitner shows that single-dipped negative assortative disclosure is optimal in their model. We also mention Leitner and Williams (2023), where a bank regulator discloses information about the design of a stress test to induce banks to make socially desirable investments. In this model, single-peaked disclosure is optimal.

Hence, if $\tilde{u}_y(y,x) < 0$ for all $(y,x)$ and $\int u(y,x)\mathrm{d}\mu = 0$, then

$$\int u_y(y,x)\mathrm{d}\mu(x) = g(y)\int \tilde{u}_y(y,x)\mathrm{d}\mu(x) + g'(y)\int \tilde{u}(y,x)\mathrm{d}\mu(x) = g(y)\int \tilde{u}_y(y,x)\mathrm{d}\mu(x) < 0,$$

yielding Assumption 2.

$1 \implies 3$. We rely on the following lemma.

**Lemma 10.** *If Assumptions 1 and 2 hold, then there exists a continuous function $h(y)$ such that*

$$u_y(y,x) + h(y)u(y,x) < 0, \quad \textit{for all } (y,x) \in Y \times X. \tag{21}$$

Given this lemma, the required $g$ is given by

$$g(y) = e^{-\int_0^y h(\tilde{y})\mathrm{d}\tilde{y}},$$

as follows from

$$\tilde{u}_y(y,x) = \frac{\partial}{\partial y}\left(\frac{u(y,x)}{e^{-\int_0^y h(\tilde{y})\mathrm{d}\tilde{y}}}\right) = \frac{u_y(y,x) + h(y)u(y,x)}{e^{-\int_0^y h(\tilde{y})\mathrm{d}\tilde{y}}} < 0.$$

*Proof of Lemma 10.* Fix $y \in Y$. Let $M_+(X)$ be the set of positive Borel measures on $X$. Define the set $C \subset \mathbb{R}^3$ as follows

$$C = \left\{\left(\int u(y,x)\mathrm{d}\mu(x), \int u_y(y,x)\mathrm{d}\mu(x) - z, \int \mathrm{d}\mu(x)\right) \mid \mu \in M_+(X),\ z \geq 0\right\}.$$

Clearly, $C$ is a convex cone.

Moreover, $C$ is closed, because $u(y,x)$ and $u_y(y,x)$ are continuous in $x$. To see this, let sequences $\mu_n \in M_+(X)$ and $z_n \in \mathbb{R}^n_+$ be such that

$$\int u(y,x)\mathrm{d}\mu_n(x) \to c_1, \int u_y(y,x)\mathrm{d}\mu_n(x) - z_n \to c_2, \int \mathrm{d}\mu_n(x) \to c_3$$

for some $(c_1, c_2, c_3) \in \mathbb{R}^3$. It follows from $\int \mathrm{d}\mu_n(x) \to c_3$ that all $\mu_n$ belong to a compact subset of positive measures whose total variation is bounded by $\sup_n \int \mathrm{d}\mu_n(x)$, and hence, up to extraction of a subsequence, $\mu_n \to \mu \in M_+(X)$, with $\int \mathrm{d}\mu(x) = c_3$. Since $u(y,x)$ and $u_y(y,x)$ are continuous in $x$, we get $\int u(y,x)\mathrm{d}\mu_n(x) \to \int u(y,x)\mathrm{d}\mu(x) = c_1$ and $\int u_y(y,x)\mathrm{d}\mu_n(x) \to \int u_y(y,x)\mathrm{d}\mu(x)$. Hence, $z_n \to \int u_y(y,x)\mathrm{d}\mu(x) - c_2 = z \geq 0$. In sum,

$$\int u(y,x)\mathrm{d}\mu(x) = c_1, \int u_y(y,x)\mathrm{d}\mu(x) - z = c_2, \int \mathrm{d}\mu(x) = c_3,$$

showing that $C$ is closed.

Next, notice that Assumption 2 implies that $(0,0,1) \notin C$. Thus, by the separation theorem (e.g., Corollary 5.84 in Aliprantis and Border 2006), there exists $\beta \in \mathbb{R}^3$ such that, for all $\mu \in M_+(X)$ and $z \geq 0$,

$$0\beta_1 + 0\beta_2 + 1\beta_3 < 0 \leq \left(\int u(y,x)\mathrm{d}\mu(x)\right)\beta_1 + \left(\int u_y(y,x)\mathrm{d}\mu(x) - z\right)\beta_2 + \left(\int \mathrm{d}\mu(x)\right)\beta_3,$$

or equivalently

$$\begin{aligned} u(y,x)\beta_1 + u_y(y,x)\beta_2 + \beta_3 &\geq 0, \quad \text{for all } x \in X, \\ -\beta_2 &\geq 0, \\ \beta_3 &< 0. \end{aligned} \tag{22}$$

We now show that there exists a scalar $h(y) \in \mathbb{R}$ satisfying

$$u_y(y,x) + h(y)u(y,x) < 0, \quad \text{for all } x \in X. \tag{23}$$

There are two cases. First, if $\beta_2 < 0$ then $h(y) = \beta_1/\beta_2 \in \mathbb{R}$ satisfies (23). Second, if $\beta_2 = 0$ then (22) implies that

$$u(y,x)\beta_1 \geq -\beta_3 > 0, \quad \text{for all } x \in X.$$

Thus, we have either (i) $u(y,x) > 0$ for all $x \in X$, so, taking into account continuity of $u(y,x)$ and $u_y(y,x)$ in $x$,

$$h(y) = \min_{x \in X}\left\{-\frac{u_y(y,x)}{u(y,x)}\right\} - 1 \in \mathbb{R}$$

satisfies (23); or (ii) $u(y,x) < 0$ for all $x \in X$, so

$$h(y) = \max_{x \in X}\left\{-\frac{u_y(y,x)}{u(y,x)}\right\} + 1 \in \mathbb{R}$$

satisfies (23).

It remains to show that if for all $y \in Y$ there exists $h(y) \in \mathbb{R}$ satisfying (23), then there exists a continuous function $\tilde{h} : Y \to \mathbb{R}$ satisfying (23). Define a correspondence $\varphi : Y \rightrightarrows \mathbb{R}$,

$$\varphi(y) = \{r \in \mathbb{R} : u_y(y,x) + ru(y,x) < 0, \quad \text{for all } x \in X\}.$$

Note that $\varphi$ is nonempty valued by assumption, and is clearly convex valued. In addition, $\varphi$ has open lower sections, because for each $r \in \mathbb{R}$ the set

$$\{y \in Y : u_y(y,x) + ru(y,x) < 0, \quad \text{for all } x \in X\}$$

is open, since $u_y$ and $u$ are continuous on the compact set $Y \times X$. Thus, by Browder's Selection Theorem (Theorem 17.63 in Aliprantis and Border 2006), $\varphi$ admits a continuous selection $\tilde{h}$, which by construction satisfies (23). □

F.2. **Proof of Lemma 1.** The proof of Lemma 1 remains valid without Assumption 4 and when $X$ is an arbitrary compact metric space.

By Corollary 2 in Dworczak and Kolotilin (2024), it suffices to show that $W$ is Lipschitz on $\Delta(X)$, endowed with the Kantorovich-Rubinstein distance

$$d_{KR}(\mu, \eta) = \sup \left\{ \int_X p(x) \mathrm{d}(\mu - \eta)(x) : p \text{ is 1-Lipschitz on } X \right\}, \quad \text{for all } \mu, \eta \in \Delta(X).$$

Recall that the Kantorovich-Rubinstein distance metrizes the weak* topology on $\Delta(X)$ (e.g., Theorem 6.9 in Villani 2009).

Let $L_{V_y}$, $L_{V_x}$, and $L_{u_x}$ be the maximum values of $|V_y|$, $|V_x|$, and $|u_x|$ on $Y \times [0,1]$, which are well-defined because $V_y$, $V_x$, and $u_x$ are continuous, by Assumption 1, and $Y \times [0,1]$ is compact. Then $V(y,x)$ is $L_{V_y}$-Lipschitz in $y$ for all $x$, $V(y,x)$ is $L_{V_x}$-Lipschitz in $x$ for all $y$, and $u(y,x)$ is $L_{u_x}$-Lipschitz in $x$ for all $y$. Moreover, let $l_{u_y}$ be the minimum value of $-\int_X u_y(\gamma(\mu), x) \mathrm{d}\mu(x)$ on $\Delta(X)$, which is well-defined by Assumption 1 and is strictly positive (i.e., $l_{u_y} > 0$), by Assumption 2. Note that $\gamma$ is $L_{u_x}/l_{u_y}$-Lipschitz on $\Delta(X)$, because, by the implicit function theorem, the derivative of $\gamma(\mu + \rho(\eta - \mu))$ with respect to $\rho$ at any $\rho \in [0,1]$ and $\mu, \eta \in \Delta(X)$ satisfies

$$\begin{aligned} \left| \frac{\mathrm{d}}{\mathrm{d}\rho} \gamma(\mu + \rho(\eta - \mu)) \right| &= \left| \frac{\int u(\gamma(\mu + \rho(\eta - \mu)), x) \mathrm{d}(\eta - \mu)(x)}{-\int u_y(\gamma(\mu + \rho(\eta - \mu)), x) \mathrm{d}(\mu + \rho(\eta - \mu))(x)} \right| \\ &\leq \left| \frac{1}{l_{u_y}} \int u(\gamma(\mu + \rho(\eta - \mu)), x) \mathrm{d}(\eta - \mu)(x) \right| \\ &\leq \frac{L_{u_x}}{l_{u_y}} d_{KR}(\eta, \mu), \end{aligned}$$

where the last inequality holds by the definition of $d_{KR}$ and $L_{u_x}$-Lipschitz continuity of $u(y,x)$ in $x$ for all $y$. Now, for any $\mu, \eta \in \Delta(X)$, we have

$$\begin{aligned} |W(\eta) - W(\mu)| &= \left| \int (V(\gamma(\eta), x) - V(\gamma(\mu), x)) \mathrm{d}\eta(x) + \int V(\gamma(\mu), x) \mathrm{d}(\eta - \mu)(x) \right| \\ &\leq \int |V(\gamma(\eta), x) - V(\gamma(\mu), x))| \, \mathrm{d}\eta(x) + \left| \int V(\gamma(\mu), x) \mathrm{d}(\eta - \mu)(x)) \right| \\ &\leq L_{V_y} \frac{L_{u_x}}{l_{u_y}} d_{KR}(\eta, \mu) + L_{V_x} d_{KR}(\eta, \mu), \end{aligned}$$

showing that $W$ is Lipschitz on $\Delta(X)$.

F.3. **Proof of Remark 1.** The proof of Remark 1 remains valid if Assumption 4 is replaced with strict single-crossing of $u(y, x)$ in $x$.

Recall that to define $\Lambda$ we took an arbitrary solution $p$ to (D). Also, recall (18) stating that

$$p(x) = V(\gamma(\mu), x) + q(\gamma(\mu))u(\gamma(\mu), x), \quad \text{for all } \mu \in \Lambda \text{ and } x \in \operatorname{supp}(\mu).$$

Fix any solution $\tau$ to (P), so $\operatorname{supp}(\tau) \subset \Lambda$. Let $X^\star = \cup_{\mu \in \operatorname{supp}(\tau)} \operatorname{supp}(\mu)$. Then, by (BP), we have $\phi(X^\star) = 1$, so the closure of $X^\star$ is $X$.

Next, take any $x \in X^\star$. If there is $\mu \in \operatorname{supp}(\tau)$ and $x \in \operatorname{supp}(\mu)$ such that $\gamma(\delta_x) = \gamma(\mu)$, then $p(x) = V(\gamma(\delta_x), x)$. Otherwise, there is $\mu \in \operatorname{supp}(\tau)$ and $x, x' \in \operatorname{supp}(\tau)$ such that either $x < \chi(\gamma(\mu)) < x'$ or $x' < \chi(\gamma(\mu)) < x$. Suppose that $x < \chi(\gamma(\mu)) < x'$ (the other case is analogous and omitted). By Theorem 1, we have

$$\begin{aligned} V_y(\gamma(\mu), x) + q(\gamma(\mu))u_y(\gamma(\mu), x) + q'(\gamma(\mu))u(\gamma(\mu), x) = 0, \\ V_y(\gamma(\mu), x') + q(\gamma(\mu))u_y(\gamma(\mu), x') + q'(\gamma(\mu))u(\gamma(\mu), x') = 0. \end{aligned}$$

Adding the first equation multiplied by $u(\gamma(\mu), x')$ and the second multiplied by $-u(\gamma(\mu), x)$, we obtain

$$q(\gamma(\mu)) = -\frac{v(\gamma(\mu), x)u(\gamma(\mu), x') - v(\gamma(\mu), x')u(\gamma(\mu), x)}{u_y(\gamma(\mu), x)u(\gamma(\mu), x') - u_y(\gamma(\mu), x')u(\gamma(\mu), x)},$$

which is well-defined because the denominator is strictly negative by Assumption 2. Consequently, $p(x) = V(\gamma(\mu), x) + q(\gamma(\mu))u(\gamma(\mu), x)$. In sum, for each $x \in X^\star$, an arbitrary solution $p$ to (D) is determined by a fixed solution $\tau$ to (P). Moreover, since $X$ is the closure of $X^\star$, there is a unique continuous extension of $p$ from $X^\star$ to $X$. This shows that there is a unique $p \in L(X)$ that solves (D).

F.4. **Proof of Theorem 2.** We first prove the second part where the twist condition holds. The proof of this part remains valid if Assumption 4 is replaced with strict single-crossing of $u(y, x)$ in $x$.

Suppose by contradiction that there exists $\mu \in \Lambda$ with $x_1 < x_2 < x_3$ in $\operatorname{supp}(\mu)$. Then, by the definition of $\gamma(\mu)$ and strict single-crossing of $u(y, x)$ in $x$, we have $\min \operatorname{supp}(\mu) < \chi(\gamma(\mu)) < \max \operatorname{supp}(\mu)$. Thus, by redefining $x_1 = \min \operatorname{supp}(\mu)$ and $x_3 = \max \operatorname{supp}(\mu)$ if necessary, we can assume that $x_1 < \chi(\gamma(\mu)) < x_3$. So, by the twist condition, the rows of the matrix $S$ are linearly independent, which contradicts the fact that (4) holds at $(\gamma(\mu), x_1)$, $(\gamma(\mu), x_2)$, and $(\gamma(\mu), x_3)$. Thus, $|\operatorname{supp}(\mu)| \leq 2$ for all $\mu \in \Lambda$, implying that every optimal signal is pairwise.

We now turn to the first part. The proof of this part does not require Assumption 4, and it remains valid when $X$ is an arbitrary compact metric space.

For any $\mu \in \Delta(X)$, denote the set of distributions of posteriors with average posterior equal to $\mu$ by

$$\Delta_2(\mu) = \left\{ \tau \in \Delta(\Delta(X)) : \int_{\Delta(X)} \eta \mathrm{d}\tau(\eta) = \mu \right\}.$$

Let $\Delta_2^{Bin}(\mu) \subset \Delta_2(\mu)$ denote the set of such distributions where in addition the posterior is always supported on at most two states:

$$\Delta_2^{Bin}(\mu) = \left\{ \tau \in \Delta_2(\mu) : \mathrm{supp}(\tau) \subset \Delta_1^{Bin} \right\},$$

where

$$\Delta_1^{Bin} = \left\{ \eta \in \Delta(X) : |\mathrm{supp}(\eta)| \leq 2 \right\}.$$

We wish to show that for each $\tau \in \Delta_2(\phi)$, there exists $\hat{\tau} \in \Delta_2^{Bin}(\phi)$ such that $\pi_{\hat{\tau}} = \pi_\tau$.

We set the stage by defining some key objects and establishing their properties. Define $\Delta_1 = \Delta(X)$ and $\Delta_2 = \Delta(\Delta(X))$. Since $X$ is compact, the sets $\Delta_1$ and $\Delta_2$ are also compact (in the weak* topology), by Prokhorov's Theorem (Theorem 15.11 in Aliprantis and Border 2006). In addition, $\Delta_2(\mu)$ is compact, since it is a closed subset of the compact set $\Delta_2$.

Define the correspondence $P : \Delta_1 \rightrightarrows \Delta_1$ as

$$P(\mu) = \left\{ \eta \in \Delta_1 : \int_X u(\gamma(\mu), x) \mathrm{d}\eta(x) = 0 \right\}.$$

For each $\mu \in \Delta_1$, $P(\mu)$ is a *moment set*—a set of probability measures $\eta \in \Delta_1$ satisfying a given moment condition (e.g., Winkler 1988). By Assumption 2, we have, for all $\mu, \eta \in \Delta_1$,

$$\eta \in P(\mu) \iff \gamma(\mu) = \gamma(\eta). \tag{24}$$

Clearly, $P(\mu)$ is nonempty (as $\mu \in P(\mu)$) and convex. Since $u(y, x)$ is continuous in $x$, $P(\mu)$ is a closed subset of $\Delta_1$, and hence is compact. Also, the correspondence $P$ has a closed graph. Indeed, consider two sequences $\mu_n \to \mu \in \Delta_1$ and $\eta_n \to \eta \in \Delta_1$ with $\mu_n \in \Delta_1$ and $\eta_n \in P(\mu_n)$, so that

$$\int_X u(\gamma(\mu_n), x) \mathrm{d}\eta_n(x) = 0.$$

Note that $\gamma(\mu)$ is a continuous function of $\mu$, by Berge's theorem (Theorem 17.31 in Aliprantis and Border 2006). Since $u$ is also continuous, by Corollary 15.7 in

Aliprantis and Border (2006), we have

$$\int_X u(\gamma(\mu), x)\,\mathrm{d}\eta(x) = 0,$$

proving that $\eta \in P(\mu)$, so $P$ has a closed graph.

Define the correspondence $E : \Delta_1 \rightrightarrows \Delta_1$ as

$$E(\mu) = P(\mu) \cap \Delta_1^{Bin} = \{\eta \in P(\mu) : |\operatorname{supp}\eta| \leq 2\}.$$

Notice that for each $\mu \in \Delta_1$, the support of $\mu$ is well-defined, by Theorem 12.14 in Aliprantis and Border (2006). Moreover, from the proof of Theorem 15.8 in Aliprantis and Border (2006), it follows that $\Delta_1^{Bin}$ is a closed subset of $\Delta_1$, so both $\Delta_1^{Bin}$ and $E(\mu)$ are compact.

Define the correspondence $\Sigma : \Delta_1 \rightrightarrows \Delta_2$ as

$$\Sigma(\mu) = \left\{\sigma \in \Delta(E(\mu)) : \mu = \int_{E(\mu)} \eta\,\mathrm{d}\sigma(\eta)\right\}.$$

Lemma 12 shows that the correspondence $\Sigma$ admits a measurable selection. In turn, Lemma 12 relies on Lemma 11.

**Lemma 11.** *Let Assumptions 1 and 2 hold. For any $y \in Y$ and $\mu \in \Delta(X)$ such that $\int u(y,x)\mathrm{d}\mu(x) = 0$, there exists $\sigma_\mu \in \Delta(\Delta(X))$ such that $\int \eta \mathrm{d}\sigma_\mu(\eta) = \mu$ and for each $\eta \in \operatorname{supp}(\sigma_\mu)$ we have $\int u(y,x)\mathrm{d}\eta(x) = 0$ and $|\operatorname{supp}(\eta)| \leq 2$.*

*Proof.* Follows immediately from the Choquet Theorem (Theorem 3.1 in Winkler 1988) and Richter-Rogosinsky's Theorem (Theorem 2.1 in Winkler 1988) □

**Lemma 12.** *There exists a measurable function $\mu \mapsto \sigma_\mu \in \Sigma(\mu)$.*

*Proof.* The correspondence $\Sigma$ is nonempty-valued, by Lemma 11. Next, fix $\mu \in \Delta_1$, and consider a sequence $\sigma_n \to \Sigma \in \Delta_2$ with $\sigma_n \in \Sigma(\mu)$. By the Portmanteau Theorem (Theorem 15.3 in Aliprantis and Border 2006), we have

$$\int_{E(\mu)} \eta\,\mathrm{d}\sigma_n(\eta) \to \int_{E(\mu)} \eta\,\mathrm{d}\sigma(\eta) \quad \text{and} \quad \limsup_n \sigma_n(E(\mu)) \leq \sigma(E(\mu)),$$

where the last inequality holds because $E(\mu)$ is closed. Thus,

$$\int_{E(\mu)} \eta\,\mathrm{d}\sigma(\eta) = \mu \quad \text{and} \quad 1 = \limsup_n \sigma_n(E(\mu)) \leq \sigma(E(\mu)) \leq 1,$$

proving that $\sigma \in \Sigma(\mu)$. Thus, $\Sigma$ is closed-valued.

Next, consider two sequences $\mu_n \to \mu \in \Delta_1$ and $\sigma_n \to \sigma \in \Delta_2$ with $\mu_n \in \Delta_1$ and $\sigma_n \in \Sigma(\mu_n)$, so that

$$\mu_n = \int_{E(\mu)} \eta \mathrm{d}\sigma_n(\eta), \quad \sigma_n\left(\Delta_1^{Bin}\right) = 1, \quad \text{and} \quad \sigma_n\left(P\left(\mu_n\right)\right) = 1.$$

The Portmanteau Theorem implies that $\mu = \int \eta \mathrm{d}\sigma(\eta)$ and $\sigma\left(\Delta_1^{Bin}\right) = 1$, since $\Delta_1^{Bin}$ is closed. Define $\overline{P}(\mu_n)$ as the closure of $\cup_{k=n}^{\infty} P(\mu_k)$. By construction, $P(\mu_k) \subset \overline{P}(\mu_k) \subset \overline{P}(\mu_n)$ for $k \geq n$, so the Portmanteau Theorem implies that $\sigma(\overline{P}(\mu_n)) = 1$. Moreover, $\overline{P}(\mu_n) \downarrow \overline{P} \subset P(\mu)$, because $P$ has a closed graph. Hence, $\sigma(P(\mu)) = 1$, by the continuity of probability measures (Theorem 10.8 in Aliprantis and Border 2006). That is, $\sigma \in \Sigma(\mu)$, showing that the correspondence $\Sigma$ has a closed graph.

Therefore, $\Sigma$ is measurable, by Theorem 18.20 in Aliprantis and Border (2006), as well as nonempty- and closed-valued. Hence, there exists a measurable function $\mu \mapsto \sigma_\mu \in \Sigma(\mu)$, by Theorem 18.13 in Aliprantis and Border (2006). □

Finally, taking a measurable selection, for each $\tau \in \Delta_2(\phi)$, define $\hat{\tau} \in \Delta_2$ as

$$\hat{\tau}\left(\widetilde{\Delta}_1\right) = \int_{\Delta_1} \sigma_\mu\left(\widetilde{\Delta}_1\right) \mathrm{d}\tau(\mu) \tag{25}$$

for every measurable set $\widetilde{\Delta}_1 \subset \Delta_1$. By construction, $\hat{\tau} \in \Delta_2^{Bin}(\phi)$, since

$$\hat{\tau}(\Delta_1^{Bin}) = \int_{\Delta_1} \sigma_\mu(\Delta_1^{Bin}) \mathrm{d}\tau(\mu) = 1$$

and

$$\phi = \int_{\Delta_1} \mu \mathrm{d}\tau(\mu) = \int_{\Delta_1} \left(\int_{E(\mu)} \eta \mathrm{d}\sigma_\mu(\eta)\right) \mathrm{d}\tau(\mu) = \int_{\Delta_1} \eta \mathrm{d}\hat{\tau}(\eta),$$

where the first equality holds by $\tau \in \Delta_2(\phi)$, the second by $\sigma_\mu \in \Sigma$, and the third by (25). Similarly, for each measurable $\widetilde{Y} \subset Y$ and $\widetilde{X} \subset X$, we have

$$\begin{aligned}
\pi_\tau(\widetilde{Y}, \widetilde{X}) &= \int_{\Delta_1} 1\{\gamma(\mu) \in \widetilde{Y}\} \mu(\widetilde{X}) \mathrm{d}\tau(\mu) \\
&= \int_{\Delta_1} 1\{\gamma(\mu) \in \widetilde{Y}\} \left(\int_{E(\mu)} \eta(\widetilde{X}) \mathrm{d}\sigma_\mu(\eta)\right) \mathrm{d}\tau(\mu) \\
&= \int_{\Delta_1} \left(\int_{E(\mu)} 1\{\gamma(\eta) \in \widetilde{Y}\} \eta(\widetilde{X}) \mathrm{d}\sigma_\mu(\eta)\right) \mathrm{d}\tau(\mu) \\
&= \int_{\Delta_1} 1\{\gamma(\eta) \in \widetilde{Y}\} \eta(\widetilde{X}) \mathrm{d}\hat{\tau}(\eta) \\
&= \pi_{\hat{\tau}}(\widetilde{Y}, \widetilde{X}),
\end{aligned}$$

where the second equality holds by $\sigma_\mu \in \Sigma$, the third by (24) and $E(\mu) \subset P(\mu)$, and the fourth by (25).

F.5. **Proof of Corollary 1.** The proof of Corollary 1 remains valid if Assumption 4 is replaced with strict single-crossing of $u(y,x)$ in $x$. Since $|X| \geq 3$, strict single-crossing of $u(y,x)$ in $x$ implies that there exist $x_1 < x_2$ in $X$ such that $x_1 < \chi(\gamma(\phi)) < x_2$.

Suppose that no disclosure is optimal. Then, by part 3 of Theorem 1, it follows that (4) holds for $\mu = \phi$ and all $x \in X$, so there exist constants $q(\gamma(\phi)), q'(\gamma(\phi)) \in \mathbb{R}$ such that

$$V_y(\gamma(\phi), x) = -q(\gamma(\phi))u_y(\gamma(\phi), x) - q'(\gamma(\phi))u(\gamma(\phi), x) \quad \text{for all } x \in X.$$

Thus, $V_y(\gamma(\phi), \cdot)$ lies in a linear space $L$ spanned by $u_y(\gamma(\phi), \cdot)$ and $u(\gamma(\phi), \cdot)$, whose dimension is at most 2. But a generic $V_y(\gamma(\phi), \cdot)$ lies in a linear space whose dimension is at least 3, since $|X| \geq 3$, and thus it does not belong to $L$, showing that generically no disclosure is suboptimal.

Finally, suppose that no disclosure is optimal for all priors. It suffices to show that $W$ is given by (6) for some functions $m$, $l$, and $H$, as then Corollary 1 in Kolotilin, Mylovanov, and Zapechelnyuk (2022) implies that no disclosure is optimal for all priors iff $H$ is concave. By Assumption 3 and part 3 of Theorem 1, there exist functions $q_0, q_1$ such that

$$V_y(y,x) + q_0(y)u_y(y,x) + q_1(y)u(y,x) = 0, \quad \text{for all } (y,x) \in [0,1]^2. \tag{26}$$

First, consider the linear receiver case. Note that (26) simplifies to

$$V_y(y,x) = q_0(y) + q_1(y)(y-x).$$

Thus,

$$V(y,x) = \int_0^y (q_0(\tilde{y}) + q_1(\tilde{y})(\tilde{y} - x))\mathrm{d}\tilde{y} + V(0,x),$$

and

$$W(\mu) = \int_0^{\mathbb{E}_\mu[x]} (q_0(\tilde{y}) + q_1(\tilde{y})(\tilde{y} - \mathbb{E}_\mu[x]))\mathrm{d}\tilde{y} + \mathbb{E}_\mu[V(0,x)].$$

So (6) holds with $m(x) = x$, $l(x) = V(0,x)$, and $H(y) = \int_0^y (q_0(\tilde{y}) + q_1(\tilde{y})(\tilde{y} - y))\mathrm{d}\tilde{y}$.

Second, consider the state-independent sender case. We have $q_0(y) > 0$ for all $y \in [0,1]$, by (3) and Assumptions 2 and 4. Differentiating (26) with respect to $x$ yields

$$u_{yx}(y,x) = -\frac{q_1(y)}{q_0(y)}u_x(y,x).$$

Thus,

$$u(y,x) = (u(0,x) - u(0,0))e^{-\int_0^y \frac{q_1(\tilde{y})}{q_0(\tilde{y})}\mathrm{d}\tilde{y}} + u(y,0),$$

implying that $\gamma(\mu) = a(\mathbb{E}_\mu[u(0,x)])$ for some function $a$. Hence

$$W(\mu) = V(a(\mathbb{E}_\mu[u(0,x)])).$$

So (6) holds with $m(x) = u(0,x)$, $l(x) = 0$, and $H(y) = V(a(y))$.

F.6. **Proof of Remark 2.** Let $\Lambda$ be strictly single-dipped. Note that there do not exist distinct $\mu$ and $\eta$ in $\Lambda$ such that $\gamma(\mu) = \gamma(\eta)$, as otherwise $\mu/2 + \eta/2$, with $|\operatorname{supp}(\mu/2 + \eta/2)| \geq 3$, would also be in $\Lambda$, contradicting that $\Lambda$ is pairwise. Thus, there exist unique functions $\chi_1$ and $\chi_2$ from $Y_\Lambda$ to $X$ such that $\operatorname{supp}(\mu) = \{\chi_1(\gamma(\mu)), \chi_1(\gamma(\mu))\}$ and $\chi_1(\gamma(\mu)) = \chi(\gamma(\mu)) = \chi_2(\gamma(\mu))$ or $\chi_1(\gamma(\mu)) < \chi(\gamma(\mu)) < \chi_2(\gamma(\mu))$ for all $\mu \in \Lambda$. Moreover, for all $y < y'$ in $Y_\Lambda$, we have $\chi_2(y) \leq \chi_2(y')$, as otherwise there would exist $\mu, \mu' \in \Lambda$ such that $\gamma(\mu) = y$, $\gamma(\mu') = y'$, and $\chi_1(y) \leq \chi(y) < \chi(y') \leq \chi_2(y') < \chi_2(y)$ contradicting that $\Lambda$ is single-dipped. Likewise, for all $y < y'$ in $Y_\Lambda$, we have $\chi_1(y') \notin (\chi_1(y), \chi_2(y))$, as otherwise there would exist $\mu, \mu' \in \Lambda$ such that $\gamma(\mu) = y$, $\gamma(\mu') = y'$, and $\chi_1(y) < \chi_1(y') < \chi_2(y)$ contradicting that $\Lambda$ is single-dipped.

F.7. **Proof of Lemma 5.** We consider the case where $u_{yx}/u_x$ and $V_{yx}/u_x$ are increasing in $x$; the case where $u_{yx}/u_x$ and $V_{yx}/u_x$ are decreasing in $x$ is analogous and thus omitted.

Fix $x_1 < x_2 < x_3$ and $y$ such that $u(y,x_1) < 0 < u(y,x_3)$. The inequality $|S| > 0$ follows from the following displayed equations:

$$u(y,x_3) - u(y,x_1) = \int_{x_1}^{x_3} u_x(y,x)\mathrm{d}x > 0,$$

where the inequality holds by Assumption 4;

$$\begin{vmatrix} u(y,x_1) & u(y,x_3) \\ u_y(y,x_1) & u_y(y,x_3) \end{vmatrix} = -u(y,x_3)u_y(y,x_1) + u(y,x_1)u_y(y,x_3) > 0,$$

where the inequality holds by part 2 of Lemma 3;

$$\begin{vmatrix} V_y(y,x_1) & V_y(y,x_3) \\ u(y,x_1) & u(y,x_3) \end{vmatrix} = u(y,x_3)V_y(y,x_1) - u(y,x_1)V_y(y,x_3) > 0,$$

where the inequality holds by Assumption 4;

$$-\begin{vmatrix} V_y(y,x_2)-V_y(y,x_1) & V_y(y,x_3)-V_y(y,x_2) \\ u(y,x_2)-u(y,x_1) & u(y,x_3)-u(y,x_2) \end{vmatrix}$$

$$=(V_y(y,x_3)-V_y(y,x_2))(u(y,x_2)-u(y,x_1))-(V_y(y,x_2)-V_y(y,x_1))(u(y,x_3)-u(y,x_2))$$

$$=\int_{x_2}^{x_3}\int_{x_1}^{x_2}(V_{yx}(y,\tilde{x})u_x(y,x)-V_{yx}(y,x)u_x(y,\tilde{x}))\mathrm{d}x\mathrm{d}\tilde{x}\geq(>)0,$$

where the inequality holds by Assumption 4 and (strict) monotonicity of $V_{yx}/u_x$ in $x$;

$$\begin{vmatrix} u(y,x_2)-u(y,x_1) & u(y,x_3)-u(y,x_2) \\ u_y(y,x_2)-u_y(y,x_1) & u_y(y,x_3)-u_y(y,x_2) \end{vmatrix}$$

$$=(u(y,x_2)-u(y,x_1))(u_y(y,x_3)-u_y(y,x_2))-(u(y,x_3)-u(y,x_2))(u_y(y,x_2)-u_y(y,x_1))$$

$$=\int_{x_2}^{x_3}\int_{x_1}^{x_2}(u_x(y,x)u_{yx}(y,\tilde{x})-u_x(y,\tilde{x})u_{yx}(y,x))\mathrm{d}x\mathrm{d}\tilde{x}\geq(>)0,$$

where the inequality holds by Assumption 4 and (strict) monotonicity of $u_{yx}/u_x$ in $x$;

$$\frac{\begin{vmatrix} V_y(y,x_1) & V_y(y,x_2) & V_y(y,x_3) \\ u(y,x_1) & u(y,x_2) & u(y,x_3) \\ u_y(y,x_1) & u_y(y,x_2) & u_y(y,x_3) \end{vmatrix}}{\begin{vmatrix} u(y,x_1) & u(y,x_3) \\ u_y(y,x_1) & u_y(y,x_3) \end{vmatrix}}(u(y,x_3)-u(y,x_1))$$

$$=-\begin{vmatrix} V_y(y,x_2)-V_y(y,x_1) & V_y(y,x_3)-V_y(y,x_2) \\ u(y,x_2)-u(y,x_1) & u(y,x_3)-u(y,x_2) \end{vmatrix}$$

$$+\frac{\begin{vmatrix} V_y(y,x_1) & V_y(y,x_3) \\ u(y,x_1) & u(y,x_3) \end{vmatrix}}{\begin{vmatrix} u(y,x_1) & u(y,x_3) \\ u_y(y,x_1) & u_y(y,x_3) \end{vmatrix}}\begin{vmatrix} u(y,x_2)-u(y,x_1) & u(y,x_3)-u(y,x_2) \\ u_y(y,x_2)-u_y(y,x_1) & u_y(y,x_3)-u_y(y,x_2) \end{vmatrix},$$

where the equality holds by rearrangement.

F.8. **Proof of Lemma 6.** We consider the case where $u_{yx}/u_x$ and $V_{yx}/u_x$ are increasing in $x$; the case where $u_{yx}/u_x$ and $V_{yx}/u_x$ are decreasing in $x$ is analogous and thus omitted.

Fix $x_1 < x_2 < x_3$ and $y_2 > y_1$ such that $u(y_1, x_1) < 0 < u(y_1, x_2)$. The inequality $|R| > 0$ follows from the following displayed equations:

$$u(y_1, x_3) - u(y_1, x_1) = \int_{x_1}^{x_3} u_x(y_1, x)\mathrm{d}x > 0,$$

where the inequality holds by Assumption 4;

$$\begin{vmatrix} u(y_1, x_1) & u(y_1, x_3) \\ u(y_2, x_1) & u(y_2, x_3) \end{vmatrix}$$

$$= -u(y_1, x_3)u(y_2, x_1) + u(y_1, x_1)u(y_2, x_3)$$

$$= -g(y_1)\tilde{u}(y_1, x_3)g(y_2)\tilde{u}(y_2, x_1) + g(y_1)\tilde{u}(y_1, x_1)g(y_2)\tilde{u}(y_2, x_3)$$

$$= g(y_1)g(y_2)[-\tilde{u}(y_1, x_3)(\tilde{u}(y_2, x_1) - \tilde{u}(y_1, x_1)) + \tilde{u}(y_1, x_1)(\tilde{u}(y_2, x_3) - \tilde{u}(y_1, x_3))]$$

$$= g(y_1)g(y_2)\int_{y_1}^{y_2} [-\tilde{u}(y_1, x_3)\tilde{u}_y(y, x_1) + \tilde{u}(y_1, x_1)\tilde{u}_y(y, x_3)]\mathrm{d}y > 0,$$

where the inequality and the second equality hold by parts 2 and 3 of Lemma 3;

$$\begin{vmatrix} V(y_2, x_1) - V(y_1, x_1) & V(y_2, x_3) - V(y_1, x_3) \\ u(y_1, x_1) & u(y_1, x_3) \end{vmatrix}$$

$$= u(y_1, x_3)\int_{y_1}^{y_2} V_y(y, x_1)\mathrm{d}y - u(y_1, x_1)\int_{y_1}^{y_2} V_y(y, x_3)\mathrm{d}y > 0,$$

where the inequality holds by Assumption 4;

$$-\begin{vmatrix} V(y_2, x_2) - V(y_1, x_2) - V(y_2, x_1) + V(y_1, x_1) & V(y_2, x_3) - V(y_1, x_3) - V(y_2, x_2) + V(y_1, x_2) \\ u(y_1, x_2) - u(y_1, x_1) & u(y_1, x_3) - u(y_1, x_2) \end{vmatrix}$$

$$= (V(y_2, x_3) - V(y_1, x_3) - V(y_2, x_2) + V(y_1, x_2))(u(y_1, x_2) - u(y_1, x_1))$$

$$-(V(y_2, x_2) - V(y_1, x_2) - V(y_2, x_1) + V(y_1, x_1))(u(y_1, x_3) - u(y_1, x_2))$$

$$= \int_{y_1}^{y_2}\int_{x_2}^{x_3}\int_{x_1}^{x_2} (V_{yx}(y, \tilde{x})u_x(y_1, x) - V_{yx}(y, x)u_x(y_1, \tilde{x}))\mathrm{d}x\mathrm{d}\tilde{x}\mathrm{d}y \geq (>)0,$$

where the inequality holds by Assumption 4 and (strict) monotonicity of $V_{yx}/u_x$ in $x$;

$$\begin{vmatrix} u(y_1, x_2) - u(y_1, x_1) & u(y_1, x_3) - u(y_1, x_2) \\ u(y_2, x_2) - u(y_2, x_1) & u(y_2, x_3) - u(y_2, x_2) \end{vmatrix}$$

$$= (u(y_1, x_2) - u(y_1, x_1))(u(y_2, x_3) - u(y_2, x_2)) - (u(y_1, x_3) - u(y_1, x_2))(u(y_2, x_2) - u(y_2, x_1))$$

$$= \int_{x_2}^{x_3}\int_{x_1}^{x_2} (u_x(y_1, x)u_x(y_2, \tilde{x}) - u_x(y_1, \tilde{x})u_x(y_2, x))\mathrm{d}x\mathrm{d}\tilde{x} \geq (>)0,$$

where the inequality holds by Assumption 4 and (strict) monotonicity of $u_{yx}/u_x$ in $x$, which imply that, for $y_2 > y_1$ and $\tilde{x} > x$, we have

$$\ln \frac{u_x(y_1,x)u_x(y_2,\tilde{x})}{u_x(y_1,\tilde{x})u_x(y_2,x)} = \int_{y_1}^{y_2} \frac{\partial}{\partial y}[\ln u_x(y,\tilde{x}) - \ln u_x(y,x)]\mathrm{d}y = \int_{y_1}^{y_2} \left[\frac{u_{yx}(y,\tilde{x})}{u_x(y,\tilde{x})} - \frac{u_{yx}(y,x)}{u_x(y,x)}\right] \mathrm{d}y \geq (>)0;$$

$$\frac{\begin{vmatrix} V(y_2,x_1)-V(y_1,x_1) & -(V(y_2,x_2)-V(y_1,x_2)) & V(y_2,x_3)-V(y_1,x_3) \\ -u(y_1,x_1) & u(y_1,x_2) & -u(y_1,x_3) \\ u(y_2,x_1) & -u(y_2,x_2) & u(y_2,x_3) \end{vmatrix}}{\begin{vmatrix} u(y_1,x_1) & u(y_1,x_3) \\ u(y_2,x_1) & u(y_2,x_3) \end{vmatrix}}(u(y_1,x_3)-u(y_1,x_1))$$

$$= -\begin{vmatrix} V(y_2,x_2)-V(y_1,x_2)-V(y_2,x_1)+V(y_1,x_1) & V(y_2,x_3)-V(y_1,x_3)-V(y_2,x_2)+V(y_1,x_2) \\ u(y_1,x_2)-u(y_1,x_1) & u(y_1,x_3)-u(y_1,x_2) \end{vmatrix}$$

$$+\frac{\begin{vmatrix} V(y_2,x_1)-V(y_1,x_1) & V(y_2,x_3)-V(y_1,x_3) \\ u(y_1,x_1) & u(y_1,x_3) \end{vmatrix}}{\begin{vmatrix} u(y_1,x_1) & u(y_1,x_3) \\ u(y_2,x_1) & u(y_2,x_3) \end{vmatrix}} \begin{vmatrix} u(y_1,x_2)-u(y_1,x_1) & u(y_1,x_3)-u(y_1,x_2) \\ u(y_2,x_2)-u(y_2,x_1) & u(y_2,x_3)-u(y_2,x_2) \end{vmatrix},$$

where the equality holds by rearrangement.

F.9. **Proof of Lemma 7.** Fix $x_1 < x_2 < x_3$ and $y_2 > y_1$ such that $u(y_1,x_1) < 0 < u(y_1,x_3)$. The first claimed inequality follows as in the proof of Lemma 6, by Assumption 2 and $u(y_1,x_1) < 0 < u(y_1,x_3)$. We thus focus on the second and third inequalities.

As in the proof of Lemma 6, Assumption 4 and monotonicity of $u_{yx}/u_x$ in $x$ yield

$$u(y_1,x_3) > u(y_1,x_2) > u(y_1,x_1),$$
$$\frac{u(y_2,x_3)-u(y_2,x_2)}{u(y_1,x_3)-u(y_1,x_2)} \geq \frac{u(y_2,x_2)-u(y_2,x_1)}{u(y_1,x_2)-u(y_1,x_1)}.$$

There are three cases to consider.

(1) $u(y_1,x_2) = 0$. In this case, $u(y_2,x_2) < 0$, by Assumption 2. Thus,

$$u(y_2,x_2)u(y_1,x_1) > 0 = u(y_2,x_1)u(y_1,x_2),$$
$$u(y_2,x_3)u(y_1,x_2) = 0 > u(y_2,x_2)u(y_1,x_3).$$

(2) $u(y_1,x_2) > 0$. In this case, as follows from the proof of Lemma 6,

$$u(y_2,x_2)u(y_1,x_1) > u(y_2,x_1)u(y_1,x_2),$$

by Assumption 2 and $u(y_1, x_1) < 0 < u(y_1, x_2)$. Thus,

$$\frac{u(y_2, x_3) - u(y_2, x_2)}{u(y_1, x_3) - u(y_1, x_2)} \geq \frac{u(y_2, x_2) - u(y_2, x_1)}{u(y_1, x_2) - u(y_1, x_1)} > \frac{u(y_2, x_2)}{u(y_1, x_2)}$$
$$\implies u(y_2, x_3)u(y_1, x_2) > u(y_2, x_2)u(y_1, x_3).$$

(3) $u(y_1, x_2) < 0$. In this case, as follows from the proof of Lemma 6,

$$u(y_2, x_3)u(y_1, x_2) > u(y_2, x_2)u(y_1, x_3),$$

by Assumption 2 and $u(y_1, x_2) < 0 < u(y_1, x_3)$. Thus,

$$\frac{u(y_2, x_2)}{u(y_1, x_2)} > \frac{u(y_2, x_3) - u(y_2, x_2)}{u(y_1, x_3) - u(y_1, x_2)} \geq \frac{u(y_2, x_2) - u(y_2, x_1)}{u(y_1, x_2) - u(y_1, x_1)}$$
$$\implies u(y_2, x_2)u(y_1, x_1) > u(y_2, x_1)u(y_1, x_2).$$

F.10. **Proof of Lemma 8.** Fix $x_1 < x_2 < x_3$ and $y_2 < y_1$ such that $x_1 < \chi(y_1) < x_3$. As in the proof of Lemma 6, Assumption 4 and monotonicity of $V_{yx}/u_x$ in $x$ yield

$$V(y_1, x_j) - V(y_2, x_j) > 0 \quad \text{for } j = 1, 2, 3, \tag{27}$$

$$u(y_1, x_3) > u(y_1, x_2) > u(y_1, x_1), \tag{28}$$

$$\begin{aligned} &\frac{V(y_1, x_3) - V(y_2, x_3) - V(y_1, x_2) + V(y_2, x_2)}{u(y_1, x_3) - u(y_1, x_2)} \\ &\leq \frac{V(y_1, x_2) - V(y_2, x_2) - V(y_1, x_1) + V(y_2, x_1)}{u(y_1, x_2) - u(y_1, x_1)}. \end{aligned} \tag{29}$$

There are two cases to consider.

(1) $u(y_1, x_2) \geq 0$. In this case, we have

$$\frac{u(y_1, x_1)}{V(y_1, x_1) - V(y_2, x_1)} < \frac{u(y_1, x_2)}{V(y_1, x_2) - V(y_2, x_2)},$$

by (27) and $u(y_1, x_1) < 0 \leq u(y_1, x_2)$, and

$$\frac{u(y_1, x_2)}{V(y_1, x_2) - V(y_2, x_2)} < \frac{u(y_1, x_3)}{V(y_1, x_3) - V(y_2, x_3)},$$

by

$$\begin{aligned} u(y_1, x_2)(V(y_1, x_3) - V(y_2, x_3)) \leq& u(y_1, x_2)\frac{u(y_1, x_3) - u(y_1, x_1)}{u(y_1, x_2) - u(y_1, x_1)}(V(y_1, x_2) - V(y_2, x_2)) \\ <& u(y_1, x_3)(V(y_1, x_2) - V(y_2, x_2)), \end{aligned}$$

where the first inequality holds by (29), $V(y_1,x_1) > V(y_2,x_1)$, $u(y_1,x_3) > u(y_1,x_2)$, and $u(y_1,x_2) \geq 0$, and the second inequality holds by $V(y_1,x_2) > V(y_2,x_2)$, $u(y_1,x_3) > u(y_1,x_2)$, and $u(y_1,x_1) < 0$.

(2) $u(y_1,x_2) \leq 0$. In this case, we have

$$\frac{u(y_1,x_2)}{V(y_1,x_2) - V(y_2,x_2)} < \frac{u(y_1,x_3)}{V(y_1,x_3) - V(y_2,x_3)},$$

by (27) and $u(y_1,x_2) \leq 0 < u(y_1,x_3)$, and

$$\frac{u(y_1,x_1)}{V(y_1,x_1) - V(y_2,x_1)} < \frac{u(y_1,x_2)}{V(y_1,x_2) - V(y_2,x_2)},$$

by

$$\begin{aligned} -u(y_1,x_2)(V(y_1,x_1) - V(y_2,x_1)) &\leq - u(y_1,x_2)\frac{u(y_1,x_3) - u(y_1,x_1)}{u(y_1,x_3) - u(y_1,x_2)}(V(y_1,x_2) - V(y_2,x_2)) \\ &< - u(y_1,x_1)(V(y_1,x_2) - V(y_2,x_2)), \end{aligned}$$

where the first inequality holds by (29), $V(y_1,x_3) > V(y_2,x_3)$, $u(y_1,x_3) > u(y_1,x_2)$, and $u(y_1,x_2) \leq 0$, and the second inequality holds by $V(y_1,x_2) > V(y_2,x_2)$, $u(y_1,x_2) > u(y_1,x_1)$, and $u(y_1,x_3) > 0$.

F.11. **Proof of Lemma 9.** We give the proof for the single-dipped case. The proof remains valid if Assumption 4 is replaced with strict single-crossing of $u(y,x)$ in $x$. Let $\tau^n$ be any optimal signal under $V^n$, so that $\mathrm{supp}(\tau^n) \subset \Lambda^n$. Since the set of compact subsets of a compact set is compact (in the Hausdorff topology), taking a subsequence if necessary, $\Lambda^n$ converges to some compact set $\overline{\Lambda} \subset \Delta(X)$. Since the set of signals is compact (in the weak* topology), taking a subsequence if necessary, $\tau^n$ converges weakly to some signal $\tau$. Finally, since $\Lambda^n \to \overline{\Lambda}$, $\tau^n \to \tau$, and $\mathrm{supp}(\tau^n) \subset \Lambda^n$, it follows that $\mathrm{supp}(\tau) \subset \overline{\Lambda}$, by Box 1.13 in Santambrogio (2015).

We claim that $\tau$ is optimal under $V$. Since $V^n_y$ converges uniformly to $V_y$, for each $\delta > 0$ there exists $n_\delta \in \mathbb{N}$ such that, for all $n \geq n_\delta$, we have $|V^n_y(y,x) - V_y(y,x)| \leq \delta$ for all $(y,x)$. Since $\tau^n$ is optimal under $V^n$, for each signal $\tilde{\tau}$ we have

$$\begin{aligned} \int_{\Delta(X)}\int_X\int_0^y V_y(\tilde{y},x)\mathrm{d}\tilde{y}\mathrm{d}\mu(x)\mathrm{d}\tau^n(\mu) &\geq \int_{\Delta(X)}\int_X\int_0^y V^n_y(\tilde{y},x)\mathrm{d}\tilde{y}\mathrm{d}\mu(x)\mathrm{d}\tau^n(\mu) - \delta \\ &\geq \int_{\Delta(X)}\int_X\int_0^y V^n_y(\tilde{y},x)\mathrm{d}\tilde{y}\mathrm{d}\mu(x)\mathrm{d}\tilde{\tau}(\mu) - \delta \\ &\geq \int_{\Delta(X)}\int_X\int_0^y V_y(\tilde{y},x)\mathrm{d}\tilde{y}\mathrm{d}\mu(x)\mathrm{d}\tilde{\tau}(\mu) - 2\delta. \end{aligned}$$

Passing to the limit as $\delta \to 0$ and $n \to \infty$ establishes the optimality of $\tau$ under $V$.

Suppose by contradiction that $\overline{\Lambda}$ is not single-dipped. Then there exist $\mu_1, \mu_2 \in \overline{\Lambda}$ and $x_1 < x_2 < x_3$ such that $x_1, x_3 \in \operatorname{supp}(\mu_1)$, $x_2 \in \operatorname{supp}(\mu_2)$, and $\gamma(\mu_1) < \gamma(\mu_2)$. Since $\Lambda^n \to \overline{\Lambda}$, there exist $\mu_1^n, \mu_2^n \in \Lambda^n$ such that $\mu_1^n \to \mu_1$ and $\mu_2^n \to \mu_2$. Since $\gamma(\mu)$ is continuous in $\mu$ and since the support correspondence is lower hemicontinuous, by Theorem 17.14 in Aliprantis and Border (2006), it follows that there exists $n$, $\mu_1^n, \mu_2^n \in \Lambda^n$, and $x_1^n < x_2^n < x_3^n$ such that $x_1^n, x_3^n \in \operatorname{supp}(\mu_1^n)$, $x_2^n \in \operatorname{supp}(\mu_2^n)$, and $\gamma(\mu_1^n) < \gamma(\mu_2^n)$, contradicting that $\Lambda^n$ is single-dipped.

F.12. **Proof of Theorem 7.** The proof of Theorem 7 remains valid if Assumption 4 is replaced with strict single-crossing of $u(y, x)$ in $x$.

We start with the following lemma, which is also used in the proof of Theorem 6.

**Lemma 13.** *If $X = [0, 1]$ and $\Lambda$ is strictly single-dipped, then for each $y$ in $Y_\Lambda$ there exists $y' \leq y$ in $Y_\Lambda$ such that $\chi_1(y) \leq \chi_1(y') = \chi_2(y') \leq \chi_2(y)$.*

*Proof.* We prove that a required $y'$ can be constructed as

$$y' = \inf\{\tilde{y} \in Y_\Lambda : \chi_1(y) \leq \chi_1(\tilde{y}) \leq \chi_2(\tilde{y}) \leq \chi_2(y)\}.$$

By definition, $y' \leq y$. Moreover, $y' \in Y_\Lambda$, because $Y_\Lambda$ is compact. Suppose by contradiction that $\chi_1(y') < \chi_2(y')$. Let $X^\star = \cup_{\mu \in \Lambda} \operatorname{supp}(\mu)$. Since there exists an optimal signal $\tau$, which satisfies $\operatorname{supp}(\tau) \subset \Lambda$ and $\int_{\Delta(X)} \mu \mathrm{d}\tau(\mu) = \phi$, we have $\phi(X^\star) = 1$, so the closure of $X^\star$ is $X = [0, 1]$. Thus, there exists $y'' \in Y_\Lambda$ such that $\chi(y'')$ or $\chi_2(y'')$ is in $(\chi_1(y'), \chi_2(y'))$. Since $\Lambda$ is strictly single-dipped, we have $y'' < y'$ and $\chi_1(y') \leq \chi_1(y'') \leq \chi_2(y'') \leq \chi_2(y')$, contradicting the definition of $y'$. ☐

Next, we claim that if $y \in Y_\Lambda$ and $\varepsilon > 0$ are such that $\chi_1(\tilde{y}) < \chi_2(\tilde{y})$ for all $\tilde{y} \in (y - \varepsilon, y) \cap Y_\Lambda$, then $\chi_1(\tilde{y}_1) < \chi_2(\tilde{y}_2)$ for all $\tilde{y}_1, \tilde{y}_2 \in (y - \varepsilon, y) \cap Y_\Lambda$. Suppose by contradiction that there exist $\tilde{y}_1, \tilde{y}_2 \in (y - \varepsilon, y) \cap Y_\Lambda$ such that $\chi_1(\tilde{y}_1) \geq \chi_2(\tilde{y}_2)$. By Lemma 13, there exists $\tilde{y}_1' \leq \tilde{y}_1$ in $Y_\Lambda$ such that $\chi_2(\tilde{y}_1') = \chi(\tilde{y}_1') = \chi_1(\tilde{y}_1') \geq \chi_1(\tilde{y}_1) \geq \chi_2(\tilde{y}_2) \geq \chi(\tilde{y}_2)$, so $\tilde{y}_1' \in (y - \varepsilon, y) \cap Y_\Lambda$ and $\chi_1(\tilde{y}_1') = \chi_2(\tilde{y}_1')$, yielding a contradiction.

Suppose now that $\phi$ has a density. Suppose by contradiction that there are two distinct optimal signals, $\tau$ and $\tau'$. Since $\Lambda$ is strictly single-dipped, for each $y \in Y_\Lambda$, there is a unique $\mu$ in $\Lambda$ such that $\gamma(\mu) = y$, namely $\mu = \rho_y \delta_{\chi_1(y)} + (1 - \rho_y) \delta_{\chi_2(y)}$ where

$$\rho_y = \begin{cases} \frac{u(y, \chi_2(y))}{u(y, \chi_2(y)) - u(y, \chi_1(y))}, & \chi_1(y) < \chi_2(y), \\ 0, & \chi_1(y) = \chi_2(y). \end{cases}$$

Thus, distinct signals $\tau$ and $\tau'$ must induce distinct distributions $\alpha$ and $\alpha'$ over actions $\gamma(\mu)$. Let $\hat{y} = \sup\{y \in Y : \alpha([0,y]) \neq \alpha'([0,y])\} \in Y_\Lambda$, where the inclusion follows from $\alpha \neq \alpha'$ and $\alpha(Y_\Lambda) = \alpha'(Y_\Lambda) = 1$. By the regularity condition and the claim above, there exists $\varepsilon > 0$ such that either (i) $\chi_1(\tilde{y}) = \chi_2(\tilde{y})$ for all $\tilde{y} \in (\hat{y} - \varepsilon, \hat{y}) \cap Y_\Lambda$ or (ii) $\chi_1(\tilde{y}_1) < \chi_2(\tilde{y}_2)$ for all $\tilde{y}_1, \tilde{y}_2 \in (\hat{y} - \varepsilon, \hat{y}) \cap Y_\Lambda$. We will now show that $\alpha([0,\tilde{y}]) = \alpha'([0,\tilde{y}])$ for all $\tilde{y} \in (\hat{y} - \varepsilon, \hat{y})$ contradicting the definition of $\hat{y}$. Since $\chi_2$ is increasing, states $x > \chi_2(\tilde{y})$ can only induce actions $y > \tilde{y}$. Thus, since $\gamma$ is bijective from $\Lambda$ to $Y_\Lambda$ and since $\alpha([0,y]) = \alpha'([0,y])$ for all $y \geq \hat{y}$, in both cases (i) and (ii), we have, for all $\tilde{y} \in (\hat{y} - \varepsilon, \hat{y}) \cap Y_\Lambda$,

$$\begin{aligned}\phi((\chi_2(\tilde{y}), 1]) - \phi([\chi_2(\tilde{y}), 1]) &\leq \int_{[\tilde{y},\hat{y}]} (1 - \rho_y)\mathrm{d}\alpha(y) - \int_{[\tilde{y},\hat{y}]} (1 - \rho_y)\mathrm{d}\alpha'(y) \\ &\leq \phi([\chi_2(\tilde{y}), 1]) - \phi((\chi_2(\tilde{y}), 1]).\end{aligned}$$

Moreover, since $\phi$ has a density, we have $\phi((\chi_2(\tilde{y}), 1]) = \phi([\chi_2(\tilde{y}), 1])$, and hence

$$\int_{[\tilde{y},\hat{y}]} (1 - \rho_y)\mathrm{d}\alpha(y) = \int_{[\tilde{y},\hat{y}]} (1 - \rho_y)\mathrm{d}\alpha'(y).$$

Then, since $1 - \rho_y > 0$ for all $y \in Y_\Lambda$, and since $\mathrm{supp}(\alpha) \subset Y_\Lambda$ and $\mathrm{supp}(\alpha') \subset Y_\Lambda$, it follows that $\alpha([\tilde{y}, \hat{y}]) = \alpha'([\tilde{y}, \hat{y}])$ for all $\tilde{y} \in (\hat{y} - \varepsilon, \hat{y})$. Thus, since $\alpha([0,y]) = \alpha'([0,y])$ for all $y \geq \hat{y}$, it follows that $\alpha([0,\tilde{y}]) = \alpha'([0,\tilde{y}])$ for all $\tilde{y} \in (\hat{y} - \varepsilon, \hat{y})$.

F.13. **Proof of Remark 4.** The proof of Remark 4 remains valid if Assumption 4 is replaced with strict single-crossing of $u$ in $x$.

Suppose by contradiction that $\Lambda$ contains $\mu = \rho\delta_{x_1} + (1-\rho)\delta_{x_2}$, with $x_1 < x_2$ and $\rho \in (0,1)$. Denote $y = \gamma(\mu)$ and $x = \chi(y)$. By strict single-crossing of $u$ in $x$, we have $x_1 < x < x_2$. Since $X = [0,1]$ and full disclosure is optimal, we have $\delta_x \in \Lambda$. Thus,

$$\rho p(x_1) + (1-\rho)p(x_2) = \rho V(y, x_1) + (1-\rho)V(y, x_2) \quad \text{and} \quad p(x) = V(y, x).$$

Adding the two equalities gives

$$\tfrac{\rho}{2}p(x_1) + \tfrac{1}{2}p(x) + \tfrac{1-\rho}{2}p(x_2) = \tfrac{1}{2}\rho V(y, x_1) + \tfrac{1}{2}V(y, x) + \tfrac{1}{2}(1-\rho)V(y, x_2),$$

which shows that $\eta = \rho\delta_{x_1}/2 + \delta_x/2 + (1-\rho)\delta_{x_2}/2$, contradicting that $\Lambda$ is pairwise.

F.14. **Proof of Theorem 5.** The proof of Theorem 5 remains valid without Assumption 4 and when $X$ is an arbitrary compact metric space. The support of the full disclosure signal is the set of all degenerate posteriors on $X$. Thus, by Lemmas 1 and 2, full disclosure is optimal iff there exists $q \in B(Y)$ such that

$$V(\gamma(\delta_x), x) \geq V(y, x) + q(y)u(y, x), \quad \text{for all } (y, x) \in Y \times X,$$
$$\iff \frac{V(y, x_1) - V(\gamma(\delta_{x_1}), x_1)}{-u(y, x_1)} \leq q(y) \leq \frac{V(\gamma(\delta_{x_2}), x_2) - V(y, x_2)}{u(y, x_2)},$$

for all $y \in Y$ and $x_1, x_2 \in X$ such that $u(y, x_1) < 0 < u(y, x_2)$. As shown in the proof of Lemma 2, the left-hand side and right-hand side functions are bounded on $Y \times X$, so full disclosure is optimal iff, for all $y \in Y$ and $x_1, x_2 \in X$ such that $u(y, x_1) < 0 < u(y, x_2)$, we have

$$\frac{V(y, x_1) - V(\gamma(\delta_{x_1}), x_1)}{-u(y, x_1)} \leq \frac{V(\gamma(\delta_{x_2}), x_2) - V(y, x_2)}{u(y, x_2)},$$
$$\iff u(y, x_2)V(y, x_1) - u(y, x_1)V(y, x_2) \leq u(y, x_2)V(\gamma(\delta_{x_1}), x_1) - u(y, x_1)V(\gamma(\delta_{x_2}), x_2),$$
$$\iff \rho V(\gamma(\mu), x_1) + (1 - \rho)V(\gamma(\mu), x_2) \leq \rho V(\gamma(\delta_{x_1})), x_1) + (1 - \rho)V(\gamma(\delta_{x_2}), x_2),$$

where $\rho = u(y, x_2)/(u(y, x_2) - u(y, x_1))$, $\mu = \rho\delta_{x_1} + (1 - \rho)\delta_{x_2}$, and $\gamma(\mu) = y$, by the definition of $\gamma(\mu)$. To complete the proof that full disclosure is optimal iff (7) holds for all $\mu$, note that for each $y$ and $x_1, x_2 \in X$ such that $u(y, x_1) < 0 < u(y, x_2)$, we have $\rho = u(y, x_2)/(u(y, x_2) - u(y, x_1)) \in (0, 1)$; and conversely, for each $x_1 < x_2$ and $\rho \in (0, 1)$, there exists a unique $y \in (\gamma(\delta_{x_1}), \gamma(\delta_{x_2}))$ such that $\rho = u(y, x_2)/(u(y, x_2) - u(y, x_1))$.

Finally, assume that (7) holds with strict inequality for all $\mu$. Suppose by contradiction that full disclosure is not uniquely optimal. Then, by Lemmas 1 and 2, there exist $\eta \in \Lambda$ and distinct $x_1, x_2 \in \text{supp}(\eta)$. By the definition of $\gamma(\eta)$, without loss, we can assume that either $u(\gamma(\eta), x_1) = 0 = u(\gamma(\eta), x_2)$ or $u(\gamma(\eta), x_1) < 0 < u(\gamma(\eta), x_2)$. In the case $u(\gamma(\eta), x_1) = 0 = u(\gamma(\eta), x_2)$, we have $\gamma(\mu) = \gamma(\delta_{x_1}) = \gamma(\delta_{x_1})$ for $\mu = \delta_{x_1}/2 + \delta_{x_2}/2$, so

$$\tfrac{1}{2}V(\gamma(\mu), x_1) + \tfrac{1}{2}V(\gamma(\mu), x_2) = \tfrac{1}{2}V(\gamma(\delta_{x_1}), x_1) + \tfrac{1}{2}V(\gamma(\delta_{x_2}), x_2),$$

contradicting that (7) holds with strict inequality for $\mu$. In the case $u(\gamma(\eta), x_1) < 0 < u(\gamma(\eta), x_2)$, we have $\gamma(\mu) = \gamma(\eta)$ for $\mu = \rho\delta_{x_1} + (1-\rho)\delta_{x_2}$ with $\rho = u(\gamma(\eta), x_2)/(u(\gamma(\eta), x_2) - u(\gamma(\eta), x_1) \in (0, 1)$. Since $\eta \in \Lambda$ and $x_1, x_2 \in \text{supp}(\eta)$, we have

$$V(\gamma(\eta), x_1) + q(\gamma(\eta))u(\gamma(\eta), x_1) \geq V(\gamma(\delta_{x_1}), x_1),$$
$$V(\gamma(\eta), x_2) + q(\gamma(\eta))u(\gamma(\eta), x_2) \geq V(\gamma(\delta_{x_2}), x_1).$$

Adding the first inequality multiplied by $\rho$ and the second inequality multiplied by $1-\rho$ gives

$$\rho V(\gamma(\mu), x_1) + (1-\rho)V(\gamma(\mu), x_2) \geq \rho V(\gamma(\delta_{x_1}), x_1) + (1-\rho)V(\gamma(\delta_{x_2}), x_2),$$

contradicting that (7) holds with strict inequality for $\mu$.

F.15. **Proof of Corollary 2'.** Condition (7) holds because

$$\begin{aligned}
&\rho V(\rho x_1 + (1-\rho)x_2, x_1) + (1-\rho)V(\rho x_1 + (1-\rho)x_2, x_2) \\
\leq &\rho(\rho V(x_1, x_1) + (1-\rho)V(x_2, x_1)) + (1-\rho)(\rho V(x_1, x_2) + (1-\rho)V(x_2, x_2)) \\
\leq &\rho V(x_1, x_1) + (1-\rho)V(x_2, x_2),
\end{aligned}$$

where the first inequality holds because $V(y, x)$ is convex in $y$, and the second holds because $V(x_1, x_2) + V(x_2, x_1) \leq V(x_1, x_1) + V(x_2, x_2)$.

F.16. **Proof of Corollary 3.** Noting that $\rho u(\gamma(\mu), x_1) + (1-\rho)u(\gamma(\mu), x_2) = 0$ and denoting $y = \gamma(\mu)$, we infer that (12) fails if there exist $x_1 < x_2$ such that for all $y \in (\gamma(\delta_{x_1}), \gamma(\delta_{x_2}))$, we have

$$u(y, x_2)(V(y, x_1) - V(\gamma(\delta_{x_1}), x_1)) - u(y, x_1)(V(y, x_2) - V(\gamma(\delta_{x_2}), x_2)) \leq 0.$$

By Taylor's theorem and some algebra, we get

$$\begin{aligned}
&u(y, x_2)(V(y, x_1) - V(\gamma(\delta_{x_1}), x_1)) - u(y, x_1)(V(y, x_2) - V(\gamma(\delta_{x_2}), x_2)) \\
&= \frac{1}{2}u_y(y, \chi(y)) \Bigg( V_{yy}(y, \chi(y)) - \frac{V_y(y, \chi(y))u_{yy}(y, \chi(y))}{u_y(y, \chi(y))} \\
&\quad -2\frac{V_{yx}(y, \chi(y))u_y(y, \chi(y)) - V_y(y, \chi(y))u_{yx}(y, \chi(y))}{u_x(y, \chi(y))} \Bigg) \\
&\quad \cdot (y - \gamma(\delta_{x_1}))(\gamma(\delta_{x_2}) - y)(\gamma(\delta_{x_2}) - \gamma(\delta_{x_1})) \\
&\quad + o((y - \gamma(\delta_{x_1}))(\gamma(\delta_{x_2}) - y)(\gamma(\delta_{x_2}) - \gamma(\delta_{x_1}))).
\end{aligned}$$

Hence, if (13) fails at some $y$, then there exist $x_2 > x_1$ with $\gamma(\delta_{x_2}) - y > 0$ and $y - \gamma(\delta_{x_1}) > 0$ small enough such that (12) fails for all $\rho \in (0, 1)$.

Note that $\mathrm{d}\chi(y)/\mathrm{d}y = -u_y(y, \chi(y))/u_x(y, \chi(y))$, by the implicit function theorem applied to $u(y, \chi(y)) = 0$. Thus, the derivative of $q(y) = -V_y(y, \chi(y))/u_y(y, \chi(y))$ is given by

$$q'(y) = -\frac{V_{yy}(y, \chi(y))}{u_y(y, \chi(y))} + \frac{V_{yx}(y, \chi(y))}{u_x(y, \chi(y))} + \frac{V_y(y, \chi(y))u_{yy}(y, \chi(y))}{(u_y(y, \chi(y)))^2} - \frac{V_y(y, \chi(y))u_{yx}(y, \chi(y))}{u_y(y, \chi(y))u_x(y, \chi(y))}.$$

Conversely, suppose that (13), together with all other assumptions of the corollary, holds. Then, for $y > \gamma(\delta_x)$, we have

$$
\begin{aligned}
&V(y,x) - \frac{V_y(y,\chi(y))}{u_y(y,\chi(y))}u(y,x) - V(\gamma(\delta_x),x)\\
=&(V(\tilde{y},x) + q(\tilde{y})u(\tilde{y},x))|_{\gamma(\delta_x)}^{y}\\
=&\int_{\gamma(\delta_x)}^{y}[V_y(\tilde{y},x) + q(\tilde{y})u_y(\tilde{y},x) + q'(\tilde{y})u(\tilde{y},x)]\mathrm{d}\tilde{y}\\
\geq&\int_{\gamma(\delta_x)}^{y}\left[V_y(\tilde{y},x) - \frac{V_y(\tilde{y},\chi(\tilde{y}))}{u_y(\tilde{y},\chi(\tilde{y}))}u_y(\tilde{y},x)\right]\mathrm{d}\tilde{y}\\
&+\int_{\gamma(\delta_x)}^{y}\left[\frac{V_y(\tilde{y},\chi(\tilde{y}))u_{yx}(\tilde{y},\chi(\tilde{y}))}{u_y(\tilde{y},\chi(\tilde{y}))u_x(\tilde{y},\chi(\tilde{y}))} - \frac{V_{yx}(\tilde{y},\chi(\tilde{y}))}{u_x(\tilde{y},\chi(\tilde{y}))}\right]u(\tilde{y},x)\mathrm{d}\tilde{y}\\
=&\int_{\gamma(\delta_x)}^{y}\int_{x}^{\chi(\tilde{y})}\left[\frac{V_y(\tilde{y},\chi(\tilde{y}))}{u_y(\tilde{y},\chi(\tilde{y}))}u_{yx}(\tilde{y},\tilde{x}) - V_{yx}(\tilde{y},\tilde{x})\right]\mathrm{d}\tilde{x}\mathrm{d}\tilde{y}\\
&+\int_{\gamma(\delta_x)}^{y}\int_{x}^{\chi(\tilde{y})}\left[\frac{V_{yx}(\tilde{y},\chi(\tilde{y}))}{u_x(\tilde{y},\chi(\tilde{y}))} - \frac{V_y(\tilde{y},\chi(\tilde{y}))u_{yx}(\tilde{y},\chi(\tilde{y}))}{u_x(\tilde{y},\chi(\tilde{y}))}\right]u_x(\tilde{y},\tilde{x})\mathrm{d}\tilde{x}\mathrm{d}\tilde{y}\\
=&\int_{\gamma(\delta_x)}^{y}\int_{x}^{\chi(\tilde{y})}\left[\frac{V_{yx}(\tilde{y},\chi(\tilde{y}))}{u_x(\tilde{y},\chi(\tilde{y}))} - \frac{V_{yx}(\tilde{y},\tilde{x})}{u_x(\tilde{y},\tilde{x})}\right]u_x(\tilde{y},\tilde{x})\mathrm{d}\tilde{x}\mathrm{d}\tilde{y}\\
&+\int_{\gamma(\delta_x)}^{y}\int_{x}^{\chi(\tilde{y})}\frac{V_y(\tilde{y},\chi(\tilde{y}))}{-u_y(\tilde{y},\chi(\tilde{y}))}\left[\frac{u_{yx}(\tilde{y},\chi(\tilde{y}))}{u_x(\tilde{y},\chi(\tilde{y}))} - \frac{u_{yx}(\tilde{y},\tilde{x})}{u_x(\tilde{y},\tilde{x})}\right]u_x(\tilde{y},\tilde{x})\mathrm{d}\tilde{x}\mathrm{d}\tilde{y} > 0,
\end{aligned}
$$

where the first and last equalities are by rearrangement, the second and third equalities are by the fundamental theorem of calculus, the first inequality is by (13) and substitution of $q(\tilde{y})$ and $q'(\tilde{y})$, and the last inequality is by our assumptions imposed in the corollary.

By Taylor's theorem, we have, for $x_1 < x_2$ and $y \in (\gamma(\delta_{x_1}), \gamma(\delta_{x_2}))$,

$$
\begin{aligned}
&u(y,x_2)(V(y,x_1) - V(\gamma(\delta_{x_1}),x_1)) - u(y,x_1)(V(y,x_2) - V(\gamma(\delta_{x_2}),x_2))\\
&= \left[V(\gamma(\delta_{x_2}),x_1) - \frac{V_y(\gamma(\delta_{x_2}),x_2)}{u_y(\gamma(\delta_{x_2}),x_2)}u(\gamma(x_2),x_1) - V(\gamma(\delta_{x_1}),x_1)\right]\\
&\qquad\cdot(-u_y(\gamma(\delta_{x_2}),x_2))(\gamma(\delta_{x_2}) - y) + o(\gamma(\delta_{x_2}) - y).
\end{aligned}
$$

Hence (12) holds for sufficiently small $\rho > 0$.

F.17. **Proof of Proposition 1.** Define the weak order $\succsim$ on $\Delta(X)$ by $\mu \succsim \eta$ if $\tilde{\gamma}(\mu) \geq \tilde{\gamma}(\eta)$. Clearly, $\tilde{\gamma}$ satisfies Betweenness and Continuity iff $\succsim$ satisfies Axioms A1(a), A2', and A4 in Dekel (1986). Thus, by his Proposition A.1, together with the characterization (**) in Proposition 1 and the argument in Section 3.C, $\tilde{\gamma}$ satisfies Betweenness and Continuity iff there exists a continuous function $\hat{u}$ from $[0,1] \times X$ to $[0,1]$ such that $\mu \succsim \eta \iff \hat{\gamma}(\mu) \geq \hat{\gamma}(\mu)$, where $\hat{\gamma}(\mu)$, for any $\mu \in \Delta(X)$, is defined implicitly as the unique $\hat{y} \in [0,1]$ satisfying

$$\int_X \hat{u}(\hat{y}, x) \mathrm{d}\mu(x) = (<) \hat{y} \iff y = (>) \hat{y}.$$

In Dekel's construction, $\hat{\gamma}$ is continuous and thus there exists a continuous and strictly increasing function $\check{\gamma}$ such that $\hat{\gamma}(\mu) = \check{\gamma}(\tilde{\gamma}(\mu))$ for all $\mu \in \Delta(X)$. Then, $u$ defined by $u(y,x) = \hat{u}(\check{\gamma}(y), x) - \check{\gamma}(y)$ for all $(y,x)$ is as stated in the proposition.

F.18. **Proof for Example 2.** First, notice that the outcome $\pi$ that corresponds to the proposed signal is implementable. (BP') holds because, for all $y \in [\underline{y}, 1]$, the marginal distribution over actions satisfies

$$\alpha_\pi([a,1]) = \phi([0, \chi_1(y)]) + \phi([a,1]),$$

and the posterior conditional on $y$ is

$$\pi_y = \frac{\mathrm{d}\phi([0,\chi_1(y)])}{\mathrm{d}\phi([0,\chi_1(y)] + \mathrm{d}\phi([a,1])} \delta_{\chi_1(y)} + \frac{\mathrm{d}\phi([a,1])}{\mathrm{d}\phi([0,\chi_1(y)] + \mathrm{d}\phi([a,1])} \delta_{\chi_1(y)},$$

as follows from $\kappa\phi([0,\chi_1(y)]) = (1-\kappa)\phi([a,1])$, which implies that $\kappa \mathrm{d}\phi([0,\chi_1(y)]) = (1-\kappa)\mathrm{d}\phi([a,1])$ and that $\chi_1$ is a continuous, strictly decreasing function. (OB) holds because, for all $y \in [\underline{y}, 1]$,

$$\mathbb{E}_{\pi_y}[u(y,x)] = \mathbb{E}_{\pi_y}[1\{x \geq y\} - \kappa] = \pi_y([y,1]) - \kappa = 0.$$

Consider now any other implementable outcome $\tilde{\pi}$. By (OB), there exists $\tilde{\pi}_y$ with $\tilde{\pi}_y([y,1]) \geq \kappa$, as otherwise $\mathbb{E}_{\tilde{\pi}_y}[u(y,x)] < 0$. Thus, by (BP'), $\alpha_{\tilde{\pi}}([y,1]) \leq \phi([y,1])/\kappa$, as follows from

$$\phi([y,1]) = \int_Y \tilde{\pi}_{\tilde{y}}([y,1]) \mathrm{d}\alpha_{\tilde{\pi}}(\tilde{y}) \geq \int_y^1 \tilde{\pi}_{\tilde{y}}([y,1]) \mathrm{d}\alpha_{\tilde{\pi}}(\tilde{y}) \geq \kappa \alpha_{\tilde{\pi}}([y,1]).$$

Since $\alpha_\pi([y,1]) = \phi([y,1])/\kappa$, it follows that $\alpha_\pi$ first-order stochastically dominates $\alpha_{\tilde{\pi}}$, and thus, for an increasing $V$,

$$\int_{Y \times X} V(y) \mathrm{d}\pi(y,x) = \int_Y V(y) \mathrm{d}\alpha_\pi(y) \geq \int_Y V(y) \mathrm{d}\alpha_{\tilde{\pi}}(y) = \int_{Y \times X} V(y) \mathrm{d}\tilde{\pi}(y,x),$$

showing that $\pi$ is optimal.

F.19. **Proof for Example 3.** We will show that $\Lambda = \{\delta_{\chi_1(y)}/2 + \delta_{\chi_2(y)}/2 : y \in [-1,1]\}$. Then, by Theorem 7, there is a unique optimal signal. Consider a signal $\tau$

that induces the distribution over actions $\alpha$ and the only posterior inducing each action $y \in \mathrm{supp}(\alpha)$ is $\mu = \delta_{\chi_1(y)}/2 + \delta_{\chi_2(y)}/2$. By construction, $\mathrm{supp}(\tau) \subset \Lambda$. Moreover, $\int \mu \mathrm{d}\tau(\mu) = \phi$, because, for each $y \in [0,1]$,

$$\phi([\chi_2(y), 3]) = \phi([3y, 3]) = \tfrac{1}{2}\alpha([y,1]),$$
$$\phi([-1, \chi_1(-y)]) = \phi([-1,-y]) = \alpha([-1,-y]) + \tfrac{1}{2}\alpha([y,1]).$$

Hence $\tau$ is optimal. Finally, the following lemma shows that $\Lambda$ is as stated.

**Lemma 14.** *Functions*

$$p(x) = \begin{cases} T(2x), & x \in [-1,0), \\ 3T(\frac{2}{3}x), & x \in [0,3], \end{cases} \quad \textit{and} \quad q(y) = \begin{cases} \frac{2T'(2y)}{T'(0)}, & y \in [-1,0), \\ 2, & y \in [0,3]. \end{cases}$$

*satisfy* (ZP') *with equality if* $y \in [-1,1]$ *and* $x \in \{\chi_1(y), \chi_2(y)\}$, *and strict inequality otherwise.*

*Proof of Lemma 14.* Since $T$ is symmetric about 0 (i.e., $T(x-y) = -T(y-x)$) and $T'$ is strictly log-concave, it follows that $T'(0) > T'(z)$ for all $z \neq 0$ and $T(z)$ is strictly concave for $z \geq 0$. Hence, if $z_1' \leq z_1 \leq z_2 \leq z_2'$, $(z_1', z_2') \neq (z_1, z_2)$, and $\rho' z_1' + (1-\rho')z_2' = \rho z_1 + (1-\rho) z_2$, for some $z_1, z_2, z_1', z_2' \geq 0$ and $\rho, \rho' \in (0,1)$, then $\rho' T(z_1') + (1-\rho)T(z_2') < \rho T(z_1) + (1-\rho)T(z_2)$, by Jensen's inequality.

We split the analysis into six cases.

(1) For $y \in [0,3]$ and $x \in [y,3]$, (ZP') simplifies to

$$3T(\tfrac{2}{3}x) \geq T(2y) + 2T(x-y),$$

which holds with equality for $x = 3y = \chi_2(a)$ and strict inequality for $x \neq 3y$.

(2) For $y \in [0,3]$ and $x \in (0,y)$, (ZP') simplifies to

$$3T(\tfrac{2}{3}x) + 2T(y-x) \geq T(2y) + 4T(0),$$

which always holds with strict inequality.

(3) For $y \in [0,3]$ and $x \in [-1,0]$, (ZP') simplifies to

$$2T(y-x) \geq T(2y) + T(-2x),$$

which holds with equality for $x = -y = \chi_1(y)$ and strict inequality for $x \neq -y$.

(4) For $y \in [-1,0)$ and $x \in [0,3]$, (ZP') simplifies to

$$3T(\tfrac{2}{3}x) + T(-2y) \geq q(y)T(x-y) + 2T(0),$$

which always holds with strict inequality because $q(y) < 2$ and $T(x-y) > 0$.

(5) For $y \in [-1, 0)$ and $x \in (y, 0)$, (ZP') simplifies to

$$T(-2y) \geq T(-2x) + q(y)T(x-y),$$

which is equivalent to

$$\frac{T(-2y) - T(-2x)}{T'(-2y)(-2y+2x)} \geq \frac{T(x-y) - T(0)}{T'(0)(x-y)},$$

which always holds with strict inequality because $T(z)$ is strictly concave for $z \geq 0$, and thus the left-hand side is strictly greater than 1 whereas the right-hand side is strictly less than 1.

(6) For $y \in [-1, 0)$ and $x \in [-1, y]$, (ZP') simplifies to

$$T(-2y) + q(y)T(y-x) \geq T(-2x),$$

which holds with equality for $x = y = \chi_1(y)$. For $x < y$, the inequality is equivalent to

$$\frac{T(y-x) - T(0)}{T'(0)(y-x)} \geq \frac{T(-2x) - T(-2y)}{T'(-2y)(-2x+2y)},$$

which always holds with strict inequality because

$$\begin{aligned}
\frac{T(-2x) - T(-2y)}{T'(-2y)(2y-2x)} &= \frac{1}{2y-2x}\int_0^{2(y-x)} \frac{T'(z-2y)}{T'(-2y)}\mathrm{d}z \\
&< \frac{1}{2y-2x}\int_0^{2(y-x)} \frac{T'(z)}{T'(0)}\mathrm{d}z \\
&= \frac{T(2y-2x) - T(0)}{T'(0)(2y-2x)} \\
&< \frac{T(y-x) - T(0)}{T'(0)(y-x)},
\end{aligned}$$

where the first inequality holds because $T'$ is strictly log-concave, and the second inequality holds because $T(z)$ is strictly concave for $z \geq 0$. □

F.20. **Proof of Proposition 2.** Recall that most results remain valid if the condition $u_x(y, x) > 0$ in Assumption 4 is replaced with strict single-crossing of $u(y, x)$ in $x$. Clearly, $\gamma(\mu) = \mathbb{E}_\mu[x]/(1 + \mathbb{E}_\mu[x^2])$. To ensure that Assumption 3 holds, we normalize $Y = [\min_{x\in[\underline{x},\overline{x}]} \gamma(\delta_x), \max_{x\in[\underline{x},\overline{x}]} \gamma(\delta_x)]$. Assumptions 1 and 2 obviously hold. Moreover, since $\gamma(\delta_x)$ is strictly increasing on $[0, 1]$ and strictly decreasing on $[1, +\infty)$, it follows that $u(\gamma(\delta_x), x') > 0$ if $x < x' \leq 1$ and if $1 \leq x' < x$. Thus, if $\overline{x} \leq 1$, then $u(y, x)$ satisfies strict single-crossing in $x$, whereas, if $\underline{x} \geq 1$, $u(y, x)$ also satisfies strict single-crossing in $x$ once the state is redefined as $-x$. So Theorems 2, 3, 5, and 6 apply.

Lemma 15 replicates Lemma 1 and Proposition 3 in Zhang and Zhou (2016).

**Lemma 15.** *If $x_1 < x_2$ and $x_1x_2 > (<)1$, then $\rho V(\gamma(\delta_{x_1}), x_1) + (1-\rho)V(\gamma(\delta_{x_2}), x_2) > (<)\rho V(\gamma(\mu), x_1) + (1-\rho)V(\gamma(\mu), x_2)$ for all $\rho \in (0,1)$.*

*Proof.* For $\mu = \rho\delta_{x_1} + (1-\rho)\delta_{x_2}$, $\gamma(\mu) = (\rho x_1 + (1-\rho)x_2)/(1+\rho x_1^2 + (1-\rho)x_2^2)$. Thus, if $x_1 < x_2$ and $x_1x_2 > (<)1$, we have

$$\frac{\mathrm{d}}{\mathrm{d}\rho}\gamma(\mu) = \frac{(x_2-x_1)(x_1x_2-1)}{(1+\rho x_1^2+(1-\rho)x_2^2)^2} > (<)0,$$
$$\frac{\mathrm{d}^2}{\mathrm{d}\rho^2}\gamma(\mu) = \frac{(x_2-x_1)(x_1x_2-1)(x_2^2-x_1^2)}{(1+\rho x_1^2+(1-\rho)x_2^2)^3} > (<)0.$$

Define $\varphi(\rho) = \gamma(\mu)\,(\rho/x_1 + (1-\rho)/x_2)$. Thus, if $x_1 < x_2$ and $x_1x_2 > (<)1$, we have

$$\varphi''(\rho) = \left(\frac{\rho}{x_1} + \frac{1-\rho}{x_2}\right)\frac{\mathrm{d}^2}{\mathrm{d}\rho^2}\gamma(\mu) + 2\left(\frac{1}{x_1} - \frac{1}{x_2}\right)\frac{\mathrm{d}}{\mathrm{d}\rho}\gamma(\mu) > (<)0,$$

so $\varphi$ is strictly convex (concave), and $\rho\varphi(1) + (1-\rho)\varphi(0) > (<)\varphi(\rho)$. □

If $\underline{x} \geq 1$, then $x_1x_2 > 1$ for all $\underline{x}_1 \leq x_1 < x_2$, so full disclosure is uniquely optimal by Theorem 5 and Lemma 15. Assume henceforth that $\underline{x} \leq 1$.

After some algebra, we get, for all $y$ and $x_1 < x_2 < x_3$,

$$|S| = \frac{(x_3-x_2)(x_3-x_1)(x_2-x_1)(1-x_2x_3-x_1x_3-x_1x_2)}{x_1x_2x_3}$$

If $\overline{x} \leq 1/\sqrt{3}$ ($\underline{x} \geq 1/\sqrt{3}$), then $|S| > (<)0$ for all $x_1 < x_2 < x_3 \leq \overline{x}$ ($\underline{x} \leq x_1 < x_2 < x_3$), so $\Lambda$ is pairwise by Theorem 2. Proposition 4 in Zhang and Zhou (2016) derives a version of this result for a finite set $X$.

Moreover, if $\overline{x} \leq 1/\sqrt{3}$ ($\underline{x} \geq 1/\sqrt{3}$), then $\Lambda$ is single-dipped (-peaked), as follows from Theorem 3 with

$$\beta = \begin{pmatrix} u(y_2,x_3)u(y_1,x_2) - u(y_2,x_2)u(y_1,x_3) \\ u(y_2,x_3)u(y_1,x_1) - u(y_2,x_1)u(y_1,x_3) \\ u(y_2,x_2)u(y_1,x_1) - u(y_2,x_1)u(y_1,x_2) \end{pmatrix}$$
$$\left(\beta = -\begin{pmatrix} u(y_2,x_3)u(y_1,x_2) - u(y_2,x_2)u(y_1,x_3) \\ u(y_2,x_3)u(y_1,x_1) - u(y_2,x_1)u(y_1,x_3) \\ u(y_2,x_2)u(y_1,x_1) - u(y_2,x_1)u(y_1,x_2) \end{pmatrix}\right),$$

because, for $y < y'$ and $x < x'$ with $xx' < 1$, we have

$$u(y', x')u(y, x) - u(y', x)u(y, x') = (y' - y)(x' - x)(1 - xx') > 0,$$

and

$$R\beta = \begin{pmatrix} (y_2 - y_1)^2|S| \\ 0 \\ 0 \end{pmatrix} \gneq \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \left( R\beta = \begin{pmatrix} -(y_2 - y_1)^2|S| \\ 0 \\ 0 \end{pmatrix} \gneq \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right).$$

Thus $\Lambda$ is strictly single-dipped (-peaked) if $\overline{x} \leq 1/\sqrt{3}$ ($\underline{x} \geq 1/\sqrt{3}$). Finally, since, by Lemma 15, (12) holds for all $\rho \in (0, 1)$, Theorem 6 yields that, if $\overline{x} \leq 1/\sqrt{3}$ ($\underline{x} \geq 1/\sqrt{3}$), then the optimal signal is unique and single-dipped (-peaked) negative assortative.

F.21. **Proof of Proposition 3.** Suppose by contradiction that an optimal outcome assigns positive probability to a strictly single-dipped triple $(y_1, x_1)$, $(y_2, x_2)$, $(y_1, x_3)$, with $x_1 < x_2 < x_3$, $y_2 < y_1$, and $x_1 \leq x_0 \leq x_3$. Consider a perturbation that reallocates mass $\beta_1\varepsilon$ on $x_1$ and mass $\beta_3\varepsilon$ on $x_3$ from $y_1$ to $y_2$, while reallocating mass $\beta_2\varepsilon$ on $x_2$ from $y_2$ to $y_1$ where $\varepsilon > 0$ is small enough and $\beta = (\beta_1, \beta_2, \beta_3)$ is given by

$$\beta = \begin{cases} \left(0, \frac{1}{(x_2 - x_0)g(y_2|x_2)}, \frac{1}{(x_2 - x_0)g(y_2|x_3)}\right), & x_2 > x_0, \\ (0, 1, 0), & x_2 = x_0, \\ \left(\frac{1}{(x_0 - x_1)g(y_1|x_1)}, \frac{1}{(x_0 - x_2)g(y_1|x_2)}, 0\right), & x_2 < x_0, \end{cases}$$

where $x_1 < x_2 < x_3$, $y_2 < y_1$, and $x_1 \leq x_0 \leq x_3$. We focus on the case $x_0 < x_2$, as the other cases are analogous. The above perturbation increases action $y_1$, because, by strict log-submodularity of $g$,

$$u(y_1, x_2)y_2 - u(y_1, x_3)y_3 = \frac{g(y_1|x_2)}{g(y_2|x_2)} - \frac{g(y_1|x_3)}{g(y_2|x_3)} > 0.$$

Moreover, the same perturbation also increases the sender's expected utility for fixed $y_1, y_2$. This follows because

$$\begin{aligned} &(V(y_1, x_2) - V(y_2, x_2))y_2 - (V(y_1, x_3) - V(y_2, x_3))y_3 \\ =& \left(\frac{G(y_1|x_2) - G(y_2|x_2)}{(x_2 - x_0)g(y_2|x_2)} - \frac{G(y_1|x_3) - G(y_2|x_3)}{(x_3 - x_0)g(y_2|x_3)}\right) \\ >& \frac{1}{(x_2 - x_0)} \left(\frac{G(y_1|x_2) - G(y_2|x_2)}{g(y_2|x_2)} - \frac{G(y_1|x_3) - G(y_2|x_3)}{g(y_2|x_3)}\right) \end{aligned}$$

$$= \frac{1}{(x_2 - x_0)} \int_{y_2}^{y_1} \left( \frac{g(t|x_2)}{g(y_2|x_2)} - \frac{g(t|x_3)}{g(y_2|x_3)} \right) \mathrm{d}t \geq 0,$$

where the first inequality is by $x_0 < x_2 < x_3$ and the second inequality is by log-submodularity of $g$. Thus, this perturbation is strictly profitable for the sender, so every optimal outcome is single-peaked.

F.22. **Proof of Proposition 4.** As shown by Kamenica and Gentzkow (2011), there exists an optimal outcome with a finite support. Suppose the support contains a strictly single-peaked triple $(y_1, x_1)$, $(y_2, x_2)$, $(y_1, x_3)$, with $x_1 < x_2 < x_3$, $y_1 < a_2$, and $x_1 < a_1 < x_3$. Notice that $V(y_1, x_3) \neq -\infty$ (so $y_1 \geq \sigma(x_3)$), as otherwise the sender's expected utility would be $-\infty$, which cannot be optimal. Taking into account that $\sigma(x) = x$ for $x \leq x_0$ gives $y_1 > x_0$. Thus, the first row in $R$ is zero. Consider a perturbation that shifts weights $\beta_1 = (x_3 - x_2)\varepsilon$ and $\beta_3 = (x_2 - x_1)\varepsilon$ on $x_1$ and $x_3$ from $y_1$ to $y_2$ and shifts weight $\beta_2 = (x_3 - x_1)\varepsilon$ from $y_2$ to $y_1$, where $\varepsilon$ takes the maximum value such that $\beta_1 \leq \pi(\{(y_1, x_1\})$, $\beta_2 \leq \pi(\{(y_2, x_2\})$, $\beta_3 \leq \pi(\{(y_1, x_3\})$, so that a strictly single-peaked triple is removed. This perturbation holds fixed $y_1$ and $y_2$ and thus does not change the sender's expected utility, since the first row in $R$ is zero. Repeating such perturbations until all strictly single-peaked triples are removed (a finite number of times since $\mathrm{supp}(\pi)$ is finite) yields a single-dipped outcome that is weakly preferred by the sender.

## References


Aliprantis, C. D., and K. Border (2006): *Infinite Dimensional Analysis: A Hitchhiker's Guide*. New York: Springer.

Dekel, E. (1986): "An Axiomatic Characterization of Preferences under Uncertainty: Weakening the Independence Axiom," *Journal of Economic Theory*, 40(2), 304–318.

Dworczak, P., and A. Kolotilin (2024): "The Persuasion Duality," *Theoretical Economics*, 19(4), 1701–1755.

Garcia, D., and M. Tsur (2021): "Information Design in Competitive Insurance Markets," *Journal of Economic Theory*, 191, 105–160.

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.

INOSTROZA, N., AND A. PAVAN (2025): "Adversarial Coordination and Public Information Design," *Theoretical Economics*, forthcoming.

KAMENICA, E., AND M. GENTZKOW (2011): "Bayesian Persuasion," *American Economic Review*, 101, 2590–2615.

KOLOTILIN, A., T. MYLOVANOV, AND A. ZAPECHELNYUK (2022): "Censorship as Optimal Persuasion," *Theoretical Economics*, 17(2), 561–585.

LEITNER, Y., AND B. WILLIAMS (2023): "Model Secrecy and Stress Tests," *Journal of Finance*, 78(2), 1055–1095.

SANTAMBROGIO, F. (2015): *Optimal Transport for Applied Mathematicians*, vol. 55. Springer.

VILLANI, C. (2009): *Optimal Transport: Old and New*, vol. 338. Berlin: Springer.

WINKLER, G. (1988): "Extreme Points of Moment Sets," *Mathematics of Operations Research*, 13(4), 581–587.

ZHANG, J., AND J. ZHOU (2016): "Information Disclosure in Contests: A Bayesian Persuasion Approach," *Economic Journal*, 126(597), 2197–2217.