Online Appendix

OA.1 Interpretation of duality

In this appendix, we interpret the persuasion problem as a linear production problem of Gale (1960). The states are economic resources, and the probability measure μ_0 is a producer's endowment of resources. The set $\Delta(\Omega)$ is the set of linear production processes available to the producer. A process $\mu \in \Delta(\Omega)$ operated at unit level consumes the measure μ of resources and generates income $V(\mu)$. A production plan τ describes the level at which each process μ is operated. The primal problem is for the producer to find a production plan that exhausts the endowment μ_0 and maximizes the total income.

To interpret the dual problem, imagine that there is a wholesaler who wants to buy out the producer. The wholesaler sets a unit price $p(\omega)$ for each resource ω. The producer's (opportunity) cost of operating a process μ at unit level is thus $\int_\Omega p(\omega)\mathrm{d}\mu(\omega)$. A price function p is feasible for the wholesaler if the income generated by each process of the producer is not greater than the cost of operating the process, which makes the producer willing to sell all the resources. The dual problem is for the wholesaler to find feasible prices that minimize the total cost of buying up all the resources.[36]

Weak duality then states that the total income generated by the producer cannot exceed the total cost of the resources under feasible prices, which make the producer willing to sell the resources. Thus, if there exists a plan for the producer and feasible prices for the wholesaler that equalize the total income with the total cost, then this plan must be optimal for the producer, and the prices must be optimal for the wholesaler. Moreover, complementary slackness states that a feasible plan and supporting prices are optimal if and only if the cost of each operated process is equal to the income it generates. Finally, strong duality states that there exists a feasible plan for the producer and feasible prices for the wholesaler such that the cost of each operated process is equal to the income it generates.

To see why optimal prices must be convex in moment persuasion, note that a measure $\mu \in \Delta(\Omega)$ of resources and one unit of resource $x = \mathbb{E}_\mu[\omega]$ are now equivalent for the producer. If prices failed to be convex, the producer could sell at effectively higher prices by engaging in such "mean-preserving" transformations of the resources. Thus, the wholesaler offers convex prices to begin with.

---

[36] A similar interpretation of states as resources in the context of persuasion appears in Galperti and Perego (2018). Dworczak and Martini (2019) offer an interpretation with the producer replaced by a consumer, production plans by consumption bundles, and the wholesaler by a Walrasian auctioneer who sets prices in a "Persuasion economy" to clear the market.

OA.2 Conditions for optimality of full disclosure and no disclosure

In this appendix, we illustrate our general duality results of Section 3, by studying conditions for optimality of two extreme information structures: full disclosure (distribution $\tau_F \in \mathcal{T}(\mu_0)$ uniquely characterized by attaching probability one to the set of Dirac probability measures on Ω) and no disclosure (distribution $\tau_N \in \mathcal{T}(\mu_0)$ that attaches probability one to the prior μ_0). We argue that strong duality makes the well-known sufficient conditions necessary.

Suppose that μ_0 has full support on Ω and let V be Lipschitz on $\Delta(\Omega)$ so that, by Theorems 3 and 4, dual attainment holds. Then, full disclosure τ_F is optimal if and only if V lies below a linear function that passes through each extreme point $(\delta_\omega, V(\delta_\omega))$:

$$V(\mu) \leq \int_\Omega V(\delta_\omega) \mathrm{d}\mu(\omega) \text{ for all } \mu \in \Delta(\Omega). \tag{F}$$

No disclosure τ_N is optimal if and only if

$$V \text{ is superdifferentiable at } \mu_0. \tag{N}$$

To prove these two observations, note that Theorem 3 implies that the dual problem (D) has an optimal solution. Thus, by Corollary 1, a feasible distribution $\tau \in \mathcal{T}(\mu_0)$ is optimal if and only if the optimal price function $p \in \mathcal{P}(V)$ satisfies (C). The support of τ_F is the set of all Dirac probability measures δ_ω on Ω, so (C) simplifies to $p(\omega) = V(\delta_\omega)$ for all $\omega \in \Omega$. Thus, τ_F is optimal if and only if $V(\delta_\omega)$, treated as a function of ω, belongs to $\mathcal{P}(V)$—this simplifies to (F). Similarly, the condition for optimality of τ_N follows from the observation that feasibility of p along with (C) is equivalent to p being the supergradient of V at the prior, yielding (N).

Because sufficiency follows from weak duality (Theorem 1), the above conditions are sufficient even if we relax the assumptions on V and μ_0. However, we show that when dual attainment fails, these conditions are no longer necessary.

Let μ_0 be the Lebesgue measure on $\Omega = [0,1]$ and let $V(\mu) = \mathbf{1}_{\{\mu = \delta_0/2 + \delta_1/2\}}$. Since $\mu_0(\{0,1\}) = 0$, there does not exist $\tau \in \mathcal{T}(\mu_0)$ with $\tau(\delta_0/2 + \delta_1/2) > 0$, so each feasible distribution $\tau \in \mathcal{T}(\mu_0)$ is optimal. However, the conditions for optimality of τ_F and τ_N both fail. In particular, (F) does not hold at $\mu = \delta_0/2 + \delta_1/2$. As for condition (N), suppose that $p \in \mathcal{P}(V)$ is its supergradient. Then, we would need $\int_0^1 p(\omega)\mathrm{d}\omega = 0$, so that the supporting hyperplane defined by p touches V at the prior. But since p is Lipschitz and non-negative, this implies that p is identically 0; hence, the hyperplane defined by p does not lie above the graph of V at $\delta_0/2 + \delta_1/2$.

The above arguments indirectly show that the dual problem does not have an optimal

solution. Indeed, the dual problem is to find a non-negative Lipschitz function p satisfying $p(0)/2 + p(1)/2 \geq 1$ that minimizes $\int_0^1 p(\omega)\mathrm{d}\omega$. We know from Theorem 2, that the infimum is 0. Clearly, the infimum is not attained: It is approximated by a sequence of Lipschitz functions that take value 1 at $\omega = 0$ and $\omega = 1$, and converge to 0 on $(0, 1)$.

OA.3 Two duality formulations for moment persuasion

In this appendix, we complement the analysis of Section 4 by formulating the dual problem for moment persuasion. We also introduce an alternative formulation of the dual problem, and show that the price function from Theorem 5 solves both of these problems. This in turn allows us to sharpen the connection between our results and existing duality methods in the next section.

The problem dual to $(\mathrm{P_M})$ is to find functions $p : X \to \mathbb{R}$ and $q : X \to \mathbb{R}^N$ to

$$\text{minimize} \int_\Omega p(\omega)\mathrm{d}\mu_0(\omega)$$
$$\text{subject to } p(y) \geq v(x) + q(x) \cdot (y - x) \text{ for all } x, y \in X, \tag{$\mathrm{D_M}$}$$
$$p \text{ is Lipschitz on } X, q \text{ is measurable and bounded on } X.$$

This duality formulation is a consequence of the fact that in our primal problem we represent feasible solutions as joint distributions of moments and states (similarly to Kolotilin, 2018 and Kolotilin et al., 2023). The dual variable p is a multiplier on the Bayes-plausibility constraint, while the dual variable q is a multiplier on the martingale constraint.

When, instead, feasibility for the primal problem is described in terms of marginal distributions of moments using a mean-preserving spread constraint (as in Dworczak and Martini, 2019 and Dizdar and Kováč, 2020), we can write the dual problem as finding a function $p : X \to \mathbb{R}$ to

$$\text{minimize} \int_\Omega p(\omega)\mathrm{d}\mu_0(\omega)$$
$$\text{subject to } p(x) \geq v(x) \text{ for all } x \in X, \tag{$\mathrm{D'_M}$}$$
$$p \text{ is convex and Lipschitz on } X.$$

In both dual formulations, only prices for states in Ω matter. However, as Theorem 5 formally shows, we can always extend these prices to the space of moments X, which provides additional insights about the structure of the solution to the primal problem $(\mathrm{P_M})$.

We now show that these problems can both be treated as duals to $(\mathrm{P_M})$ in the sense that their values provide the relevant upper bound on the value of $(\mathrm{P_M})$ that is tight and attained by the price function identified in Theorem 5.

Proposition OA.1.

1. **Weak duality:** *If v is measurable and bounded, then for any π feasible for $(\mathrm{P_M})$ and any p feasible for either $(\mathrm{D_M})$ or $(\mathrm{D'_M})$, $\int_X v(x)\mathrm{d}\pi_X(x) \leq \int_\Omega p(\omega)\mathrm{d}\mu_0(\omega)$.*

2. **No duality gap and primal attainment:** *If v is bounded and upper semi-continuous, then there exists an optimal solution to $(\mathrm{P_M})$, and the problems $(\mathrm{P_M})$, $(\mathrm{D_M})$, $(\mathrm{D'_M})$ all have the same value.*

3. **Dual attainment:** *If v is Lipschitz, then the price function $\bar{p}$ from Theorem 5 solves $(\mathrm{D'_M})$, and together with the function q from condition 2 of Theorem 5 solves $(\mathrm{D_M})$.*

Proof. <u>Weak duality.</u> Suppose that π is feasible for $(\mathrm{P_M})$. If (p, q) is feasible for $(\mathrm{D_M})$, then

$$\int_X v(x)\mathrm{d}\pi_X(x) = \int_{X\times\Omega}(v(x)+q(x)\cdot(\omega-x))\mathrm{d}\pi(x,\,\omega) \leq \int_{X\times\Omega}p(\omega)\mathrm{d}\pi(x,\,\omega) = \int_\Omega p(\omega)\mathrm{d}\mu_0(\omega).$$

If instead p is feasible for $(\mathrm{D'_M})$, then

$$\int_X v(x)\mathrm{d}\pi_X(x) \leq \int_X p(x)\mathrm{d}\pi_X(x) \leq \int_\Omega p(\omega)\mathrm{d}\mu_0(\omega).$$

<u>No duality gap and primal attainment.</u> When v is bounded and upper semi-continuous on X, the corresponding V is also bounded and upper semi-continuous on $\Delta(\Omega)$, and hence, by Lemma 2, the problem $(\mathrm{P_M})$ has an optimal solution $\pi^\star \in \Pi(\mu_0)$.

Thus, weak duality above implies that $\max(\mathrm{P_M}) \leq \inf(\mathrm{D_M})$. Moreover, if p is feasible for $(\mathrm{D'_M})$, then, by Corollary 13.3.3 in Rockafellar (1970), p has a bounded subgradient (which we denote q), so that, for all $x, y \in X$,

$$p(y) \geq p(x) + q(x)\cdot(y-x) \geq v(x) + q(x)\cdot(y-x),$$

showing that $(p,\, q)$ is feasible for $(\mathrm{D_M})$ and hence $\max(\mathrm{P_M}) \leq \inf(\mathrm{D_M}) \leq \inf(\mathrm{D'_M})$.

Thus, it suffices to show that $\max(\mathrm{P_M}) = \inf(\mathrm{D'_M})$. The proof is essentially the same as the proof of Lemma 5. Let $\mathcal{P}_M(v)$ denote the sets of functions $p : X \to \mathbb{R}$ feasible for $(\mathrm{D'_M})$. By Baire's Theorem, there exists a non-increasing sequence of Lipschitz functions v_k converging pointwise to v. Let $\pi_k^\star$ denote an optimal solution to $(\mathrm{P_M})$ with the objective function v_k. For each $k \in \mathbb{N}$, we have

$$\int_{X\times\Omega} v(x)\mathrm{d}\pi^\star(x,\omega) \leq \inf_{p\in\mathcal{P}_M(v)}\int_\Omega p(\omega)\mathrm{d}\mu_0(\omega) \leq \min_{p\in\mathcal{P}_M(v_k)}\int_\Omega p(\omega)\mathrm{d}\mu_0(\omega) = \int_{X\times\Omega} v_k(x)\mathrm{d}\pi_k^\star(x,\omega),$$

4

where the first inequality holds by $\max{(\mathrm{P_M})} \leq \inf{(\mathrm{D'_M})}$, the second inequality holds by $\mathcal{P}_M(v_k) \subset \mathcal{P}_M(v)$ for $v_k \geq v$, and the equality holds by Theorem 5. It is thus sufficient to show that

$$\lim_{k\to\infty} \int_{X\times\Omega} v_k(x)\mathrm{d}\pi_k^\star(x,\omega) \leq \int_{X\times\Omega} v(x)\mathrm{d}\pi^\star(x,\omega).$$

Thanks to compactness of $\Pi(\mu_0)$, up to extraction of a subsequence, we can suppose that $\pi_k^\star$ converges weakly to some $\overline{\pi} \in \Pi(\mu_0)$. Then for each $j \in \mathbb{N}$, we have

$$\lim_{k\to\infty} \int_{X\times\Omega} v_k(x)\mathrm{d}\pi_k^\star(x,\omega) \leq \lim_{k\to\infty} \int_{X\times\Omega} v_j(x)\mathrm{d}\pi_k^\star(x,\omega) = \int_{X\times\Omega} v_j(x)\mathrm{d}\overline{\pi}(x,\omega),$$

where the first inequality holds because $v_k \leq v_j$ for $k \geq j$, and the equality holds because v_j is (Lipschitz) continuous and $\pi_k^\star \to \overline{\pi}$. Then letting j go to infinity and invoking the monotone convergence theorem,

$$\lim_{j\to\infty} \int_{X\times\Omega} v_j(x)\mathrm{d}\overline{\pi}(x,\omega) = \int_{X\times\Omega} v(x)\mathrm{d}\overline{\pi}(x,\omega),$$

we obtain

$$\lim_{k\to\infty} \int_{X\times\Omega} v_k(x)\mathrm{d}\pi_k^\star(x,\omega) \leq \int_{X\times\Omega} v(x)\mathrm{d}\overline{\pi}(x,\omega) \leq \int_{X\times\Omega} v(x)\mathrm{d}\pi^\star(x,\omega),$$

where the last inequality holds because $\pi^\star$ is an optimal solution to $(\mathrm{P_M})$. This establishes that $\max{(\mathrm{P_M})} = \inf{(\mathrm{D_M})} = \inf{(\mathrm{D'_M})}$.

Dual attainment. When v is Lipschitz, Theorem 5 guarantees existence of $\bar{p}$ and q with all required properties, and such that for any π optimal for $(\mathrm{P_M})$,

$$\int_X v(x)\mathrm{d}\pi_X(x) = \int_\Omega \bar{p}(\omega)\mathrm{d}\mu_0(\omega).$$

It follows that $\bar{p}$ solves $(\mathrm{D'_M})$ and $(\bar{p},\, q)$ solve $(\mathrm{D_M})$. $\qquad\square$

Proposition OA.1 formalizes our claim from Section 4 that the two conditions in Theorem 5 correspond to two alternative formulations of the problem dual to $(\mathrm{P_M})$. At the same time, the proposition shows that these two problems have the same solution, at least under the conditions of Theorem 5. This observation allows us to describe the exact connection between our general duality result and existing duality approaches to moment persuasions.

OA.4 Detailed relationship to other duality methods

The one-dimensional moment persuasion problem has received special attention (see, for example, Gentzkow and Kamenica, 2016, Kolotilin et al., 2017, Kolotilin, 2018, Dworczak and Martini, 2019, and Dizdar and Kováč, 2020). When the objective function is Lipschitz, Theorem 5 generalizes Theorems 1 and 2 in Dworczak and Martini (2019): By a simple transformation, condition 1 of Theorem 5 establishes existence of a convex and (Lipschitz) continuous function $p^\star$ and a cumulative distribution function $G^\star$ of moments (a mean-preserving contraction of F_0) such that

$$p^\star \geq v,$$
$$\text{supp}(G^\star) \subseteq \{x \in X : p^\star(x) = v(x)\},$$
$$\int_\Omega p^\star(x)\mathrm{d}F_0(x) = \int_\Omega p^\star(x)\mathrm{d}G^\star(x).$$

Moreover, the theorem resolves (positively) the conjecture of Dworczak and Martini that if the objective function V is measurable with respect to a moment $m(\omega)$, then so is the corresponding price function.

It is worth noting that we impose stronger regularity conditions on the price function compared to Dworczak and Martini. In our dual formulation (D'_M), we assume that prices p are Lipschitz continuous, while Dworczak and Martini only assume continuity. The general trade-off is that stronger regularity conditions on the dual variable make it more difficult to prove that the dual problem has a solution in the assumed class, but—conditional on proving existence—impose tighter structure on the solutions to the primal problem. We impose a stronger condition on p because Lipschitz continuity is directly implied by Theorem 5. Dizdar and Kováč (2020) prove, under weaker assumptions on the objective function, that the prices that solve the dual problem of Dworczak and Martini are in fact Lipschitz.[37] Thus, it seems that in most economically relevant cases imposing Lipschitz continuity of prices in the dual is without significant loss of generality.[38]

Kolotilin (2018), Galperti et al. (2023), and Kolotilin et al. (2023) use an alternative approach to the persuasion problem. Instead of working with an objective function V :

---

[37]Dizdar and Kováč show that the dual problem in one-dimensional moment persuasion has an optimal solution by demonstrating that feasible solutions can be restricted to a compact set of uniformly Lipschitz functions. Our proof strategy is different: We construct the optimal solution (a price function on the space of moments) from the supergradient of the concave closure of V.

[38]That being said, it is easy to come up with examples where the dual problem (D'_M) has a solution in the class of continuous functions but not in the class of Lipschitz functions. For instance, when μ_0 is fully supported on $\Omega = [0, 1]$ and $v(x) = -\sqrt{x}$, $p(x) = -\sqrt{x}$ is continuous and achieves the lower bound in (D'_M), but a Lipschitz solution does not exist.

$\Delta(\Omega) \to \mathbb{R}$, they consider a Sender and a Receiver whose utility functions are $w : A \times \Omega \to \mathbb{R}$ and $u : A \times \Omega \to \mathbb{R}$ where A is the space of the Receiver's actions. The Sender chooses a joint distribution $\pi \in \Delta(A \times \Omega)$ of the recommended action a and the state ω. On top of the Bayes plausibility constraint, π must satisfy the obedience constraint, which requires each recommended action to be incentive-compatible for the Receiver given the beliefs it induces. As noted in Section 6, it is possible to reformulate the alternative problem as our problem, and vice versa.

By setting $A = X$ in the model of Kolotilin (2018), we can draw a tight connection between the two duality approaches. For $w(a, \omega) = v(a)$ and $u(a, \omega) = -(a - \omega)^2$, the dual problem in Kolotilin (2018) is to find a continuous function $p : \Omega \to \mathbb{R}$ and a bounded measurable function $q : A \to \mathbb{R}$ to

$$\text{minimize} \ \int_\Omega p(\omega) \mathrm{d}\mu_0(\omega) \tag{$\mathrm{D_A}$}$$
$$\text{subject to } p(\omega) + q(a)(a - \omega) \geq v(a) \text{ for all } (a, \omega) \in A \times \Omega,$$

where p and q are multipliers for the Bayes plausibility and obedience constraints. Thus, the problem ($\mathrm{D_A}$) corresponds to our dual problem ($\mathrm{D_M}$), and condition 2 of Theorem 5 establishes that this problem is solved by the price function $\bar{p}$ derived from our general duality results from Section 3.

OA.5   Comments on the convex-roof construction

In this appendix, we further investigate the properties of the convex-roof construction that underlies the proof of Theorem 5. Our goal is twofold: On one hand, we are interested in regularity conditions under which the convex roof is (Lipschitz) continuous, guaranteeing that it can be used as the price function $\bar{p}$ satisfying conditions 1 and 2 of Theorem 5 (and hence as a solution to the dual problems ($\mathrm{D'_M}$) and ($\mathrm{D_M}$)). On the other hand, we show (by means of examples) that the convex roof can behave in surprisingly pathological ways when the space of moments is multi-dimensional, explaining why we need stronger assumptions to extend existing duality methods to the multi-dimensional case.

The main result in this appendix shows that if the support of the prior contains the boundary of its convex hull, then the convex roof preserves the Lipschitz constant of the objective function, and hence the convex roof could be used as a solution to problems ($\mathrm{D_M}$) and ($\mathrm{D'_M}$).

Proposition OA.2. *Let v be L-Lipschitz on X, and let Ω contain the boundary of X. Then $\check{p}$ is L-Lipschitz on X.*

Proof. By the proof of Theorem 5, there exists a price function $\bar{p} : X \to \mathbb{R}$ that is convex and L-Lipschitz. Moreover, for each $z \in X$, we have $\check{p}(z) \geq \bar{p}(z)$ and, for each $y \in \Omega$, there exists a sequence $x_n \in X$ converging to some $x \in X$ such that $q(x_n)$ converges to some $r(y) \in \mathbb{R}^N$, with $\|r(y)\| \leq L$, and

$$\check{p}(y) = \bar{p}(y) = \lim_{n \to \infty} \{v(x_n) + q(x_n) \cdot (y - x_n)\}.$$

Thus, for each $z \in X$ and each $y \in \mathrm{bd}\, X \subset \Omega$, we have

$$\check{p}(z) - \check{p}(y) \geq \lim_{n \to \infty} \{v(x_n) + q(x_n) \cdot (z - x_n) - v(x_n) - q(x_n) \cdot (y - x_n)\} = r(y) \cdot (z - y),$$

showing that $r(y)$, with $\|r(y)\| \leq L$, is a subgradient of $\check{p}$ at $y \in \mathrm{bd}\, X$.

By Theorem 7.12 in Aliprantis and Border (2006), at each $z \in \mathrm{int}(X)$, the convex roof $\check{p}$ has a subgradient $r(z) \in \mathbb{R}^N$. We claim that $\|r(z)\| \leq L$. Suppose that $r(z) \neq 0$, as otherwise the claim is trivial. Since $z \in \mathrm{int}(X)$ and $\|r(z)\| > 0$, there exists $t > 0$ such that $y := z + t r(z) \in \mathrm{bd}(X) \subset \Omega$. Hence,

$$L \|y - z\| \geq \bar{p}(y) - \bar{p}(z) = \check{p}(y) - \bar{p}(z) \geq \check{p}(y) - \check{p}(z)$$
$$\geq \check{p}(z) + r(z) \cdot (y - z) - \check{p}(z) = r(z) \cdot (y - z) = \|r(z)\| \|y - z\|,$$

showing that $\|r(z)\| \leq L$.

Thus, for each $z, y \in X$, we have

$$\check{p}(z) - \check{p}(y) \leq r(y) \cdot (z - y) \leq \|r(y)\| \|z - y\| \leq L \|y - z\|,$$

showing that $\check{p}$ is L-Lipschitz on X. $\qquad\square$

Next, we construct an example showing that the assumptions of Proposition OA.2 are not redundant: $\check{p}$ does not always preserve the Lipschitz constant of v even when $N = 2$ and Ω is finite.

Example OA.1. Let $\Omega = \{(-l, 0), (0, 1), (l, 0)\}$ with $l > 1$ and $v(x) = |x_1|$ for $x \in X$, which is 1-Lipschitz. We can apply Corollary 1 to show that full disclosure is optimal and thus p that coincides with v on Ω solves (D). Indeed, condition (C) holds, and, by Jensen's inequality,

$$V(\mu) = \left| \int_\Omega \omega_1 \mathrm{d}\mu(\omega) \right| \leq \int_\Omega |\omega_1| \mathrm{d}\mu(\omega) = \int_\Omega p(\omega) \mathrm{d}\mu(\omega) \text{ for all } \mu \in \Delta(\Omega).$$

It is easy to see that $\check{p}$ is given by $\check{p}(x) = l(1 - x_2)$ for all $x \in X$, so the Lipschitz constant of

$\check{p}$ is $l > 1$. Of course, by Theorem 5, there exists a different convex extension $\bar{p}$ of p from Ω to X (for example, consider $\bar{p} = v$ on X) that is convex, 1-Lipschitz, and satisfies $\bar{p} \geq v$. ∎

The next example demonstrates the additional difficulties that arise when the dimension of the space of moments is three (or higher). In this case, even when the objective function is continuously differentiable, and the set of extreme points of X is compact, the convex roof may be discontinuous.

Example OA.2. The example is adapted from Example 5.1 in Bucicovschi and Lebl (2013). Let

$$K = \left\{ (x_1, x_2, x_3) : \ x_1 = -1, \ x_2^2 + x_3^2 = 1 \right\} \cup \left\{ (x_1, x_2, x_3) : \ x_1 = 1, \ x_2^2 + x_3^2 = 1 \right\},$$

and $\omega^\star = (0, 0, 1)$. Define $\Omega = K \cup \{\omega^\star\}$, and note that its convex hull X is a cylinder:

$$X = \left\{ (x_1, x_2, x_3) : \ -1 \leq x_1 \leq 1, \ x_2^2 + x_3^2 \leq 1 \right\}.$$

Define the objective function as $v(x) = x_1^2$ for $x \in X$, which is Lipschitz. We can again apply Corollary 1 to show that p that coincides with v on Ω solves (D).

We will now show that the convex roof $\check{p}$ is discontinuous at $\omega^\star$. On any line segment $\{(x_1, x_2, x_3) : \ -1 \leq x_1 \leq 1, \ x_2 = y, \ x_3 = z)\}$ with $y \neq 0$ and $y^2 + z^2 = 1$, the convex roof $\check{p}$ must be identically 1. This shows that $\check{p}$ is discontinuous at $\omega^\star = (0, 0, 1)$, because $\check{p}(\omega^\star) = 0$ yet $\check{p}(\omega^n) = 1$ for the sequence $\omega^n = (0, 1/n, \sqrt{1 - 1/n^2})$ that converges to $\omega^\star$, as $n \to \infty$.

By Theorem 5, there exists a convex, Lipschitz extension $\bar{p}$ (for example, $\bar{p} = v$). ∎

Finally, we construct an instance of moment persuasion (with a discontinuous objective function) in which there exists an optimal convex and Lipschitz price function on Ω solving the original dual (D), but the price function cannot be extended to a convex and continuous function on X. This example, unlike the previous ones, goes beyond indicating a problem with the convex-roof construction; it shows that—beyond the case of a Lipschitz v— requiring the price function to be (Lipschitz) continuous on X in the multi-dimensional moment persuasion problem may be too demanding.

Example OA.3. The example is adapted from Example 5.4 in Bucicovschi and Lebl (2013). Let

$$K = \left\{ (x_1, x_2, x_3) : \ -1 \leq x_1 \leq -x_3, \ x_2^2 + x_3^2 = 1 \right\} \cup \left\{ (x_1, x_2, x_3) : \ x_1 = 1, \ x_2^2 + x_3^2 = 1 \right\},$$

and $\omega^\star = (0, 0, 1)$. Define $\Omega = K \cup \{\omega^\star\}$, and note that its convex hull X is the same cylinder:

$$X = \left\{ (x_1, x_2, x_3) : -1 \leq x_1 \leq 1, \; x_2^2 + x_3^2 \leq 1 \right\}.$$

Define the objective function

$$v(x) = \begin{cases} 1, & x \in K, \\ 0, & x = \omega^\star, \\ -1, & x \notin K \cup \{\omega^\star\}. \end{cases}$$

Because the sets K and $\{\omega^\star\}$ are closed and disjoint, the function v is upper semi-continuous.

We claim that full disclosure is optimal in this instance of moment persuasion. We can again apply Corollary 1 by defining $p = v$ on Ω. Then, p is trivially Lipschitz, and condition (C) holds, so all we have to check is that for all $x \in X$, and $\mu \in \Delta(\Omega)$ such that $\int_\Omega \omega \mathrm{d}\mu(\omega) = x$, $\int_\Omega p(\omega) \mathrm{d}\mu(\omega) \geq v(x)$. When $x \notin K$, this is trivial because $p \geq 0$. When $x \in K$, the conclusion is trivial for all μ with support in K. So the only case we have to check is when $x \in K$ but $\mathrm{supp}(\mu)$ contains the point $\omega^\star$. We will prove that this case cannot arise. Indeed, since $\omega^\star$ is an isolated point of Ω, it would have to be that $\mu(\omega^\star) > 0$ and

$$x = \mu(\omega^\star)\omega^\star + \int_K \omega \mathrm{d}\mu(\omega).$$

But $x \in K$ implies that, for almost all $\omega \in \mathrm{supp}(\mu)$, $\omega_2 = 0$ and $\omega_3 = 1$ (as otherwise $x_2^2 + x_3^2 < 1$). But the only points in K with that property are $(-1, 0, 1)$ and $(1, 0, 1)$. This is a contradiction with $\mu(\omega^\star) > 0$, because $\mu(\omega^\star) > 0$ implies that $x_1 \in (-1, 1)$.

We will now show that there does not exist a convex and continuous extension of p to X. On any line segment $\{(x_1, x_2, x_3) : -1 \leq x_1 \leq 1, x_2 = y, x_3 = z)\}$ with $y \neq 0$ and $y^2 + z^2 = 1$, the function p takes the value 1 for $x_1 \in [-1, -z] \cup \{1\}$. Hence, any convex extension $\bar{p}$ of p must be identically equal to 1 on such a line segment. This, however, means that such $\bar{p}$ must be discontinuous at $\omega^\star = (0, 0, 1)$. Indeed, $\bar{p}(\omega^\star) = 0$, but $\bar{p}(\omega^n) = 1$ for the sequence $\omega^n = (0, 1/n, \sqrt{1 - 1/n^2})$ that converges to $\omega^\star$. ∎

OA.6 An example with infinite S_x for Section 4.3.2

Suppose that μ_0 is uniformly distributed on

$$\Omega = X = \{(\omega_1, \omega_2) \in [-2, 2] \times [0, 1]\} \cup \{(\omega_1, \omega_2) \in [-2, 2] \times [-1, 0] : (\omega_1/2)^2 + \omega_2^2 \leq 1\},$$

and suppose that the objective function is

$$v(x) = \begin{cases} -(x_1^2 - 1)^2, & x_2 \geq 0, \\ -(x_1^2 + x_2^2 - 1)^2, & x_2 \leq 0. \end{cases}$$

The optimal solution to (D) is $p(x) = 0$ for all $x \in X$. Moreover,

$$S^\star = \{x \in X : x_1^2 = 1, \, x_2 \geq 0\} \cup \{x \in X : x_1^2 + x_2^2 = 1, \, x_2 \leq 0\},$$

and $\Gamma_x = \Omega$, so that $S_x = \mathrm{supp}(\pi_X)$, for all $x \in S^\star$. Thus, $\pi \in \Pi(\mu_0)$ is an optimal solution to (P_M) if and only if $S_x \subset S^\star$. Applying Jensen's inequality to a strictly concave function $\omega_2 \mapsto \sqrt{1 - \omega_2^2}$ and a strictly convex function $\omega_2 \mapsto -\sqrt{1 - \omega_2^2}$, we conclude that $\pi \in \Pi(\mu_0)$ satisfies $S_x \subset S^\star$ only if S_x contains the set $\{x \in X : x_1^2 + x_2^2 = 1, \, x_2 \leq 0\}$. That is, for $x_2 < 0$, each optimal signal must pool the states within a line segment $\{(t, x_2) \in \Omega : t > 0\}$ to induce a posterior mean $(\sqrt{1 - x_2^2}, \, x_2)$ and pool the states within a line segment $\{(t, x_2) \in \Omega : t < 0\}$ to induce a posterior mean $(-\sqrt{1 - x_2^2}, \, x_2)$. This shows that there does not exist an optimal signal such that S_x is a finite set. Since the set $\{x \in X : x_1^2 = 1, \, x_2 \geq 0\}$ consists of line segments, there are multiple optimal signals differing in how states with $\omega_2 \geq 0$ are pooled. There exists an optimal signal with $S_x = \{x \in X : x_1^2 + x_2^2 = 1, \, x_2 \leq 0\} \cup \{(-1, 1/2), (1, 1/2)\}$, so that $S_x = \mathrm{ext}(S_x)$, in line with Theorem 8, but there also exists an optimal signal with $S_x = S^\star$, so that $S_x \neq \mathrm{ext}(S_x)$.

OA.7   An explicit formula for property 1 in Proposition 2

Let g denote the density of the prior distribution μ_0 on Ω.

Lemma OA.1. *Property 1 in Proposition 2 holds if and only if for almost all $t \in [\underline{x}_1, \overline{x}_1]$,*

$$\int_{\underline{l}_t}^{\overline{l}_t} l(2f'(t) - f''(t)l)g(t + l, f(t) - f'(t)l)\mathrm{d}l = 0.$$

Proof. Define $\widetilde{\Omega} = \bigcup_{t \in [\underline{x}_1, \overline{x}_1]} I_t$ and recall that $\mu_0(\widetilde{\Omega}) = 1$. By footnote 27, $\mathbb{E}[\omega \,|\, \omega \in I_t] = (t, f(t))$ is equivalent to $\mathbb{E}[\omega_1 \,|\, \omega \in I_t] = t$. Let G be the distribution function of the posterior mean of ω_1 induced by $\pi^\star$, so that, for all $t \in [\underline{x}_1, \overline{x}_1]$, we have

$$G(t) = \int_{\bigcup_{s \in [\underline{x}_1, t]} I_s} g(\omega_1, \omega_2)\mathrm{d}\omega_1 \mathrm{d}\omega_2.$$

By the definition of the conditional expectation, property 1 in Proposition 2 holds if and

11

only if, for all $t \in [\underline{x}_1, \overline{x}_1]$, we have

$$\int_{\underline{x}_1}^t s \mathrm{d}G(s) = \int_{\bigcup_{s \in [\underline{x}_1, t]} I_s} \omega_1 g(\omega_1, \omega_2) \mathrm{d}\omega_1 \mathrm{d}\omega_2.$$

Consider a change of variables on $\widetilde{\Omega}$ given by the following transformation: $(\omega_1, \omega_2) = (t + l, f(t) - f'(t)l)$ where $t \in [\underline{x}_1, \overline{x}_1]$ and $l \in (\underline{l}_t, \overline{l}_t)$. This transformation is diffeomorphism, as (1) it is one-to-one and onto $\widetilde{\Omega}$, because $I_t \cap I_s = \emptyset$ for $t \neq s$, (2) it is continuously differentiable, because f is a twice continuously differentiable function, and (3) the Jacobian determinant is negative on $\widetilde{\Omega}$,

$$J(t, l) = \det \begin{pmatrix} \frac{\partial \omega_1}{\partial t} & \frac{\partial \omega_2}{\partial t} \\ \frac{\partial \omega_1}{\partial l} & \frac{\partial \omega_2}{\partial l} \end{pmatrix} = \det \begin{pmatrix} 1 & f'(t) - f''(t)l \\ 1 & -f'(t) \end{pmatrix} = -(2f'(t) - f''(t)l) < 0,$$

where the inequality follows from the second-order condition for the second property in Proposition 2 on the (relatively) open set I_t. Thus, by the Change of Variables Theorem (Theorem 13.49 in Aliprantis and Border, 2006 and Remark 1.3 in Villani, 2009), we have, for all $t \in [\underline{x}_1, \overline{x}_1]$,

$$G(t) = \int_{\underline{x}_1}^t \int_{\underline{l}_s}^{\overline{l}_s} |J(s, l)| g(s + l, f(s) - f'(s)l) \mathrm{d}l \mathrm{d}s$$

$$= \int_{\underline{x}_1}^t \int_{\underline{l}_s}^{\overline{l}_s} (2f'(s) - f''(s)l) g(s + l, f(s) - f'(s)l) \mathrm{d}l \mathrm{d}s,$$

and

$$\int_{\underline{x}_1}^t s \mathrm{d}G(s) = \int_{\underline{x}_1}^t \int_{\underline{l}_s}^{\overline{l}_s} (s + l) |J(s, l)| g(s + l, f(s) - f'(s)l) \mathrm{d}l \mathrm{d}s$$

$$= \int_{\underline{x}_1}^t \int_{\underline{l}_s}^{\overline{l}_s} (s + l)(2f'(s) - f''(s)l) g(s + l, f(s) - f'(s)l) \mathrm{d}l \mathrm{d}s.$$

Substituting G from the first equation to the last equation, we have, for all $t \in [\underline{x}_1, \overline{x}_1]$,

$$\int_{\underline{x}_1}^t \int_{\underline{l}_s}^{\overline{l}_s} l(2f'(s) - f''(s)l) g(s + l, f(s) - f'(s)l) \mathrm{d}l \mathrm{d}s = 0,$$

which holds if and only if the inner integral is 0 for almost all $s \in [\underline{x}_1, \overline{x}_1]$. $\qquad\square$