Extended Real-Valued Information Design

Nathan Yoder*

May 17, 2023

Abstract

This note shows that Bayesian persuasion problems still have solutions when the cost of inducing some posteriors is infinite, e.g., when information has a constant marginal cost. Consequently, the concavification approach of Kamenica and Gentzkow (2011) can be applied to these settings.

In a *Bayesian persuasion* or *information design* problem (Kamenica and Gentzkow, 2011; Kamenica, 2019), a sender chooses a Blackwell experiment on a set of states Ω to maximize the expectation of some function v of the induced posterior belief. Kamenica and Gentzkow (2011) show that this is equivalent to choosing a distribution of posterior beliefs τ whose mean is the prior $p_0 \in \Delta(\Omega)$:

$$\sup_{\tau \in \Delta(\Delta(\Omega))} \{E_\tau v(p) \text{ s.t. } E_\tau p = p_0\} \tag{1}$$

Typically, the value function v is the sender's interim payoff from a receiver's use of her experiment's result to choose a risky action. Sometimes, this payoff is modified by a function representing her cost of experimentation, as in Gentzkow and Kamenica (2014), or other payoffs from generating a certain posterior, as in Yoder (2022).

Kamenica and Gentzkow (2011) show that when v is bounded, the problem has a solution whenever v is upper semicontinuous. I show that this is also true when v is unbounded, or extended real-valued. This extends their results to settings where the sender faces an infinite cost of inducing certain posteriors. These infinite costs are natural in many settings: As Pomatto et al. (2020) point out, they are present in any application where the sender faces a constant marginal cost of experimentation.

*University of Georgia, Terry College of Business, Department of Economics; E-mail: nathan.yoder@uga.edu. This paper contains material from earlier drafts of Yoder (2022). I gratefully acknowledge the Terry College of Business' summer support through a Terry-Sanford research award.

Related literature

Among settings where some posteriors are infinitely costly to induce, the most heavily studied are those where the sender has one of the *log-likelihood ratio* (LLR) cost functions introduced by Pomatto et al. (2020). Specific cost functions in this class arise from foundations like Wald (1945) sequential sampling (Morris and Strack, 2019) and indifference to sequential learning (Bloedel and Zhong, 2021). In general, the LLR class is characterized by a cost of experimentation proportional to the Kullback-Leibler divergence between the distributions of signals conditional on different states. By formulating the information design problem as one of choosing those conditional signal distributions, Pomatto et al. (2020) observe that a solution exists when the sender has LLR costs, since Kullback-Leibler divergence is a lower semicontinuous function on the space of probability measures.

This note extends their observation beyond the LLR functional form. More generally, my results show that we need not depart from the posterior-based approach of Kamenica and Gentzkow (2011) to ensure that an extended real-valued information design problem has a solution. Instead, we can rely directly on the value function's upper semicontinuity in the induced posterior belief, just as Kamenica and Gentzkow (2011) show we can in the real-valued case.

Main results[1]

As authors such as Halac et al. (2022) and Doval and Skreta (2023) have noted, existence of a solution to (1) follows immediately from two facts:

i. If $f : S \to \mathbb{R}$ is upper semicontinuous **and bounded**, then when $\Delta(S)$ is given the weak* topology, the mapping $\tau \mapsto \int f d\tau$ is upper semicontinuous as well.[2] (e.g., Aliprantis and Border (2006) Theorem 15.5)

ii. S is compact if and only if $\Delta(S)$ is. (e.g., Aliprantis and Border (2006) Theorem 15.11)

The key to this note is showing that (i) can be extended to the case where f takes the value $-\infty$, so long as its domain is compact.

Lemma 1. *If $f : S \to \mathbb{R} \cup \{-\infty\}$ is upper semicontinuous and $S \subseteq \Delta(\Omega)$ is compact, then when $\Delta(S)$ is given the weak* topology, the mapping $\tau \mapsto \int f d\tau$ is upper semicontinuous.*

[1] Let $\overline{\mathbb{R}}$ denote the affinely extended real numbers $\mathbb{R} \cup \{\pm\infty\}$. For a set S, let $\mathrm{cl}(S)$ denote its closure; $\mathrm{conv}(S)$, its convex hull; $\dim(S)$, its dimension; $\mathrm{ri}(S)$, its relative interior; and $\Delta(S)$, the set of probability measures on S. For a function $v : S \to \overline{\mathbb{R}}$, let $\mathrm{dom}(v) \equiv \{s \in S | v(s) \in \mathbb{R}\}$ denote its effective domain; $\mathrm{Gr}(v) \equiv \{(s, v(s)) | s \in \mathrm{dom}(v)\}$, its graph; and $\mathrm{hypo}(v) \equiv \{(s, y) | s \in \mathrm{dom}(v), y \leq v(s)\}$, its hypograph. For a probability distribution τ, let $\mathrm{supp}\, \tau$ denote its support.

[2] The boundedness of v is also crucial to Kamenica and Gentzkow (2011)'s argument, which relies on the compactness of the interesting part of v's hypograph, $H = \{(p, z) \in \mathrm{hypo}(v) | z \geq \inf v(p)\}$. (See the proof of Proposition 3.1 in the online appendix to Kamenica and Gentzkow (2011).)

Proof. Suppose that for $\{\tau_n\}_{n=1}^{\infty} \subset \Delta(S)$, $\tau_n \to^{w^*} \tau$.

By (the extended real-valued version of) Baire's theorem on semicontinuous functions,[3] there exists a sequence of continuous functions $\{f_k : S \to \mathbb{R}\}_{k=1}^{\infty}$ such that $f_k \downarrow f$ pointwise. Then for each k, $\limsup_{n\to\infty} \int f d\tau_n \leq \limsup_{n\to\infty} \int f_k d\tau_n$ by monotonicity of the integral.

Moreover, since each f_k is continuous and S is compact, so is $f_k(S) \subseteq \mathbb{R}$ for each k; it follows that each f_k is bounded. Then for each k, $\lim_{n\to\infty} \int f_k d\tau_n = \int f_k d\tau$.

Finally, by Lebesgue's monotone convergence theorem, $\lim_{k\to\infty} \int f_k d\tau = \int f d\tau$. Then we have

$$\limsup_{n\to\infty} \int f d\tau_n \leq \lim_{k\to\infty} \lim_{n\to\infty} \int f_k d\tau_n = \lim_{k\to\infty} \int f_k d\tau = \int f d\tau,$$

as desired. □

Proposition 1 (Existence of Solutions)**.** *Suppose that $v : S \to \mathbb{R} \cup \{-\infty\}$ is upper semicontinuous and $S \subseteq \Delta(\Omega)$ is compact. Then $\arg\max_{\tau\in\Delta(S)}\{E_\tau v(p)$ s.t. $E_\tau p = p_0\}$ is nonempty.*

Corollary 1. *Suppose that $v : \Delta(\Omega) \to \mathbb{R} \cup \{-\infty\}$ is upper semicontinuous and Ω is compact. Then (1) has a solution.*

As it turns out, Lemma 1 also makes it possible to extend Proposition 1 to the constrained information design problem considered by Doval and Skreta (2023):

$$\max_{\tau\in\Delta(\Delta(\Omega))} \{E_\tau v(p) \text{ s.t. } E_\tau p = p_0 \text{ and } E_\tau \bar{g}_k(p) \geq 0, k = 1, \ldots, K\}. \tag{2}$$

In fact, we can allow both the objective function v and the constraint functions $\bar{g}_k$ to take the value $-\infty$.

Proposition 2. *Suppose that $v : \Delta(\Omega) \to \mathbb{R} \cup \{-\infty\}$ and $\{\bar{g}_k : \Delta(\Omega) \to \mathbb{R} \cup \{-\infty\}\}_{k=1}^{K}$ are each upper semicontinuous and that Ω is compact. Then (2) has a solution.*

Proof. Since Ω is compact, so is $\Delta(\Omega)$ (e.g., Aliprantis and Border (2006) Theorem 15.11). By Lemma 1, for each k, the mapping $\tau \mapsto \int \bar{g}_k d\tau$ is upper semicontinuous, and so $\{\tau \in \Delta(\Omega) : \int \bar{g}_k d\tau \geq 0\}$ is closed. Then by Lemma 2, $\{\tau \mid E_\tau p = p_0\} \cap \left(\bigcap_{k=1}^{K}\{\tau \in \Delta(\Omega) : \int \bar{g}_k d\tau \geq 0\}\right)$ is closed, and hence compact (since $\Delta(\Omega)$ is compact).

Then since the mapping $\tau \mapsto \int v d\tau$ is upper semicontinuous by Lemma 1, it attains a maximum on this set. □

Proposition 1 allows us to extend several of the key results from Kamenica and Gentzkow (2011) to the extended real-valued case. In particular, Proposition 3 shows that the value of the problem is given by the value function's concavification $V(p) \equiv \sup\{z \mid (p, z) \in \text{conv}(\text{hypo}(v))\}$ evaluated at the prior, and the support of the solution need not have more elements than the set of states.

[3] See, e.g., **?** Theorem 2.64.

Proposition 3 (Solutions to Extended-Real Valued Persuasion Problems)**.** *Suppose that $v : S \to \mathbb{R} \cup \{-\infty\}$ is upper semicontinuous, $S \subseteq \Delta(\Omega)$ is compact and convex, and Ω is Polish. Then for all $p_0 \in \mathrm{ri}(\mathrm{dom}\, v)$,*

i. $V(p_0) = \max_{\tau \in \Delta(S)}\{E_\tau v(p) \text{ s.t. } E_\tau p = p_0\}$.

ii. There exists $\tau^* \in \arg\max_{\tau \in \Delta(S)}\{E_\tau v(p) \text{ s.t. } E_\tau p = p_0\}$ such that $\mathrm{supp}\, \tau^*$ has no more than $\dim(S) + 1$ elements.

Note that in a persuasion problem where the set of states Ω is finite, part (ii) bounds the number of posteriors in the support of the problem's solution by $|\Omega|$, since the simplex $\Delta(\Omega)$ has dimension $|\Omega| - 1$.

Corollary 2. *Suppose that $v : \Delta(\Omega) \to \mathbb{R} \cup \{-\infty\}$ is upper semicontinuous and $|\Omega| = N$. Then for all $p_0 \in \mathrm{ri}(\mathrm{dom}\, v)$, there exists $\tau^* \in \arg\max_{\tau \in \Delta(\Omega)}\{E_\tau v(p) \text{ s.t. } E_\tau p = p_0\}$ such that $\mathrm{supp}\, \tau^*$ has no more than N elements.*

References

ALIPRANTIS, C. D. AND K. C. BORDER (2006): *Infinite Dimensional Analysis: A Hitchhiker's Guide*, Springer.

ATTOUCH, H. (1984): *Variational convergence for functions and operators*, vol. 1, Pitman Advanced Publishing Program.

BLOEDEL, A. W. AND W. ZHONG (2021): "The Cost of Optimally-Acquired Information," Working paper.

DOVAL, L. AND V. SKRETA (2023): "Constrained information design," *Mathematics of Operations Research*.

GENTZKOW, M. AND E. KAMENICA (2014): "Costly Persuasion," *American Economic Review: Papers and Proceedings*, 104, 457–462.

HALAC, M., E. LIPNOWSKI, AND D. RAPPOPORT (2022): "Addressing strategic uncertainty with incentives and information," in *AEA Papers and Proceedings*, American Economic Association, vol. 112, 431–437.

KAMENICA, E. (2019): "Bayesian Persuasion and Information Design," *Annual Review of Economics*, 11, 249–272.

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

MORRIS, S. AND P. STRACK (2019): "The Wald Problem and the Equivalence of Sequential Sampling and Ex-Ante Information Costs," *Available at SSRN 2991567*.

POMATTO, L., P. STRACK, AND O. TAMUZ (2020): "The Cost of Information," *arXiv preprint arXiv:1812.04211*.

RUBIN, H. AND O. WESLER (1958): "A note on convexity in Euclidean -space," *Proceedings of the American Mathematical Society*, 9, 522–523.

WALD, A. (1945): "Sequential Tests of Statistical Hypotheses," *The Annals of Mathematical Statistics*, 16, 117–186.

YODER, N. (2022): "Designing incentives for heterogeneous researchers," *Journal of Political Economy*, 130, 2018–2054.

Proofs[4]

Lemma 2. *If $S \subseteq \Delta(\Omega)$ is compact, the set of Bayes-plausible distributions on S, $\{\tau \in \Delta(S) \mid E_\tau p = p_0\}$, is a compact subspace of $\Delta(\Omega)$ in the weak* topology.*

Proof. Since S is compact, so is $\Delta(S)$ (e.g., Aliprantis and Border (2006) Theorem 15.11). By Lemma 1, the mapping $\tau \mapsto \int p d\tau(p)$ is both upper and lower semicontinuous, hence continuous. It follows that the preimage of $\{p_0\}$ under this mapping, $\{\tau \mid E_\tau p = p_0\}$, is closed. Then $\{\tau \mid E_\tau p = p_0\} \cap \Delta(S)$ is compact. □

Proof of Proposition 1 (Existence of Solutions) By Lemma 1, the mapping $\tau \mapsto \int v d\tau$ is upper semicontinuous on $\Delta(S)$, and so attains a maximum on any compact subspace of $\Delta(S)$ (e.g., Aliprantis and Border (2006) Theorem 2.43). The statement then follows from Lemma 2. □

Proof of Corollary 1 Since Ω is compact, so is $\Delta(\Omega)$ (e.g., Aliprantis and Border (2006) Theorem 15.11); the statement follows from Proposition 1. □

Lemma 3 (Properties of the Concavification). *Let $S \subseteq \mathbb{R}^n$ be convex and compact; let $v : S \to \mathbb{R} \cup \{-\infty\}$ be upper semicontinuous with* $\operatorname{dom} v \neq \emptyset$*; and let V be the concavification of v. Then $\bar{V}$ is bounded above;* $\operatorname{dom} V = \operatorname{conv}(\operatorname{dom} v)$*; and* $\operatorname{ri}(\operatorname{dom} v) \subseteq \operatorname{ri}(\operatorname{dom} V)$.

Proof. Concavity of V follows immediately from the definition. Since v is upper semicontinuous and S is compact, v is bounded above: for some $\bar{z} \in \mathbb{R}$, we have $v(p) \leq \bar{z}$ for all $p \in S$. Then $\{z | (p,z) \in \operatorname{conv}(\operatorname{hypo}(v))\}$ is also bounded above: if $(p,z) \in \operatorname{conv}(\operatorname{hypo}(v))$, then for some d, $\{p_i\}_{i=1}^d \subseteq S$, and $(\lambda_1, \ldots, \lambda_d) \in \Delta(\{1, \ldots, d\})$, $z \leq \sum_{i=1}^d \lambda_i v(p_i) \leq \bar{z}$. Then V is bounded above: $V(p) \equiv \sup\{z | (p,z) \in \operatorname{conv}(\operatorname{Gr}(v))\} \leq \bar{z} < \infty$ for each p.

Then $p \in \operatorname{dom} V \Leftrightarrow V(p) > -\infty$. By definition, we have $V(p) > -\infty$ if and only if there exists $(p,z) \in \operatorname{conv}(\operatorname{hypo}(v))$ with $z > -\infty$. This is true if and only if there exist d, $\{p_i\}_{i=1}^d \subseteq \operatorname{dom} v$, and $\{\lambda_i\}_{i=1}^d \subset (0,1]$ such that $\sum_{i=1}^d \lambda_i = 1$, $\sum_{i=1}^d \lambda_i v(p_i) > -\infty$, and $\sum_{i=1}^d p_i = p$. Now for any $\{\lambda_i\}_{i=1}^d \subset (0,1]$, $\sum_{i=1}^d \lambda_i v(p_i) > -\infty$ if and only if $\{p_i\}_{i=1}^d \subseteq \operatorname{dom} v$; it follows that $V(p) > -\infty \Leftrightarrow p \in \operatorname{conv}(\operatorname{dom} v)$. Hence $\operatorname{dom} V = \operatorname{conv}(\operatorname{dom} v)$. Then the relative topologies of $\operatorname{dom} v$ and $\operatorname{dom} V$ coincide, since $\operatorname{aff}(\operatorname{dom} V) = \operatorname{aff}(\operatorname{conv}(\operatorname{dom} v)) = \operatorname{aff}(\operatorname{dom} v)$; then since $\operatorname{dom} v \subseteq \operatorname{conv}(\operatorname{dom} v)$, $\operatorname{ri}(\operatorname{dom} v) \subseteq \operatorname{ri}(\operatorname{dom} V)$. □

[4]For a set S, let $\operatorname{cl}(S)$ denote its closure, and let $\operatorname{aff}(S)$ denote its affine hull.

Lemma 4. *Suppose $v : S \to \mathbb{R} \cup \{-\infty\}$ is upper semicontinuous, and $S \subseteq \Delta(\Omega)$ is convex and compact. Then v's concavification V and concave closure $\bar{V}(p) \equiv \max\{z \mid (p, z) \in \mathrm{cl}(\mathrm{conv}(\mathrm{hypo}(v)))\}$ coincide on $\mathrm{ri}(\mathrm{dom}\, v)$.*

Proof. By Lemma 3, V is bounded above. Then it is continuous on $\mathrm{ri}(\mathrm{dom}\, V)$ (e.g., Aliprantis and Border (2006) Theorem 5.43).

Then $\mathrm{hypo}(V) \cap (\mathrm{ri}(\mathrm{dom}\, V) \times \mathbb{R}) = (\mathrm{Gr}(V) \cup \mathrm{conv}(\mathrm{hypo}(v))) \cap (\mathrm{ri}(\mathrm{dom}\, V) \times \mathbb{R})$ is closed as a subset of $(\mathrm{ri}(\mathrm{dom}\, V) \times \mathbb{R})$.

Then if $p_0 \in \mathrm{ri}(\mathrm{dom}\, v)$ and (p_0, z) is a limit point of $\mathrm{conv}(\mathrm{hypo}(v))$, we must have $(p_0, z) \in \mathrm{Gr}(V) \cup \mathrm{conv}(\mathrm{hypo}(v))$ and hence $z \leq V(p_0)$. It follows from the definition of $\bar{V}$ that $\bar{V}(p) = \mathrm{conv}(p)$. $\square$

Lemma 5. *The concave closure of $v : S \to \mathbb{R} \cup \{-\infty\}$ coincides with its concave envelope: For all $p \in S$, $\bar{V}(p) = \hat{\bar{V}}(p) \equiv \inf\{f(p) \mid f \geq v$ and f is affine and continuous$\}$.*

Proof. First, we show that $\inf\{f(p) \mid f \geq v$ and f is affine and continuous$\} = \inf\{f(p) \mid f \geq v$ and f is concave and u.s.c.$\}$: By Aliprantis and Border (2006) Theorem 7.6,

$$\begin{aligned}
&\inf\{f(p) \mid f \geq v \text{ and } f \text{ is concave and u.s.c.}\}\\
&\qquad = \inf\{f(p) \mid f \text{ is affine and continuous and } f \geq g \geq v \text{ for some concave and u.s.c. } g\}\\
&\qquad = \hat{\bar{V}}(p). \quad \text{(by choosing } g = f\text{)}
\end{aligned}$$

Hence, since $\bar{V}$ is concave and u.s.c. and $\bar{V} \geq v$, we have $\bar{V}(p) \geq \hat{\bar{V}}(p)$.

Now by definition, there exists a sequence $\{(p_n, z_n)\}_{n=1}^{\infty} \subset \mathrm{conv}(\mathrm{hypo}(v))$ such that $(p_n, z_n) \to (p, \bar{V}(p))$. Then for each n, $z_n \leq \sum_{i=1}^{d} \lambda^i v(p_n^i)$ for some $\{\lambda_i, p_n^i\}_{i=1}^{d}$ such that $\sum_{i=1}^{d} \lambda^i = 1$ and $\sum_{i=1}^{d} \lambda^i p_n^i = p_n$. Then for each n, we have $\hat{\bar{V}}(p_n) = \inf\{\sum_{i=1}^{d} \lambda^i f(p_n^i) \mid f \geq v$ and f is affine and continuous$\} \geq z_n$. Then since $\hat{\bar{V}}$ is upper semicontinuous (it is the infimum of a family of upper semicontinuous functions) we have $\hat{\bar{V}}(p) \geq \limsup_{n\to\infty} \hat{\bar{V}}(p_n) \geq \lim z_n = \bar{V}(p)$. $\square$

Proof of Proposition 3 (i): By Lemmas 4 and 5 and by monotonicity of the integral, for any Bayes-plausible $\tau \in \Delta(S)$,

$$V(p_0) = \hat{\bar{V}}(p_0) = \inf\{\int_S f(p) d\tau(p) \mid f \geq v \text{ and } f \text{ is affine and continuous}\} \geq \int_S v(p) d\tau(p).$$

It follows that $V(p_0) \geq \sup_{\tau \in \Delta(S)}\{E_\tau v(p) \text{ s.t. } E_\tau p = p_0\}$.

Moreover, $V(p_0) \leq \sup_{\tau \in \Delta(S)}\{E_\tau v(p) \text{ s.t. } E_\tau p = p_0\}$: For any $(p_0, z) \in \mathrm{conv}(\mathrm{Gr}(v))$, there exist $d \in \mathbb{N}$, $\{\lambda_i\}_{i=1}^{d} \subset (0, 1]$, and $\{p_i\}_{i=1}^{d}$ such that $\sum_{i=1}^{d} \lambda_i = 1$, $\sum_{i=1}^{d} \lambda_i p_i = p_0$, and $\sum_{i=1}^{d} \lambda_i v(p_i) = z$. Then $\tau = \sum_{i=1}^{d} \lambda_i \delta_{p_i}$ is Bayes-plausible and $E_\tau v(p) = z$. The inequality then follows from the definition of V.

(ii): The statement is vacuous unless $\dim(S)$ is finite. Suppose it is. Since Ω is Polish, so is $\Delta(\Omega)$ (e.g., Aliprantis and Border (2006) Theorem 15.15). Then so is S; it follows that the topology of S is Euclidean (Aliprantis and Border (2006) Theorem 5.21). By definition, for any p, $(p, v(p)) \in \text{conv}(\text{hypo}(v))$. Then by Rubin and Wesler (1958), for any $\tau \in \Delta(S)$ with $E_\tau p = p_0$, $(p_0, E_\tau v(p)) \in \text{conv}(\text{hypo}(v))$. Then by (i), $(p_0, V(p_0)) \in \text{conv}(\text{hypo}(v))$. The result then follows from Carathéodorory's Theorem. $\square$