Divisible Atoms: Extreme and Exposed Mean-Preserving Contractions of Arbitrary Priors*

Franz Ostrizek†

October 11, 2026

Abstract

We study linear optimization over mean-preserving contractions of an arbitrary probability distribution on a compact interval. We generalize the notion of bi-pooling and define a new notion of split rigidity. For every prior,

$$\{\text{exposed}\} = \{\text{split-rigid}\} \subseteq \{\text{extreme}\} \subseteq \{\text{bi-pooling}\}.$$

Both inclusions can be strict when the prior has atoms: a bi-pooling distribution need not be extreme, and an extreme point need not be exposed. We give conditions under which the familiar equivalences are restored. We also show that censorship policies are optimal for S-shaped objectives under an arbitrary prior.

Linear optimization problems over the set of mean-preserving contractions of a given prior distribution have received considerable attention in the recent literature. In the classic formulation in information design, a designer chooses a distribution over posterior means given a prior distribution over the state to maximize an objective (Gentzkow and Kamenica, 2016). Applications include the design of grading policies (Ostrovsky and Schwarz, 2010), the disclosure of bank stress tests (Goldstein and Leitner, 2018), and media censorship (Kolotilin et al., 2022). Equivalent and closely related problems also arise outside of information design. Linear delegation admits an equivalent formulation as an optimization problem over mean-preserving contractions (Kleiner et al., 2021; Kolotilin and Zapechelnyuk, 2025). More generally, optimization under majorization constraints has been used to study the allocation of heterogeneous goods with redistributive objectives

*This paper generalizes results developed in Ostrizek (2024) for the case of S-shaped objective functions (essentially the present Proposition 7). I thank Victor Augias for comments and discussions. LLMs from Anthropic and OpenAI were used for research assistance in compiling this manuscript. All remaining errors and writing tics are my own—all dashes were written by hand.

†Department of Economics, Sciences Po, Paris. franz.ostrizek@gmail.com

(Akbarpour et al., 2024b), vaccine prioritization (Akbarpour et al., 2024a), and the allocation of positional goods (Xiao, 2026). Related formulations also connect screening and auction design to the same mathematical methods (Bergemann et al., 2026). In these settings, the constraints arise from incentive compatibility and the physical feasibility of the allocation. As is natural, this problem has also received considerable attention from a more theoretical point of view, aimed at developing solution and verification methods and characterizing general properties of the solution, sometimes under additional restrictions on the objective. Prominent contributions include the duality approach of Dworczak and Martini (2019) and the extreme-point and bi-pooling characterizations of Kleiner et al. (2021) and Arieli et al. (2023).

However, many of these results impose restrictive regularity assumptions on the primitive objects such as the designer's objective and, most notably, on the prior distribution. Typically, the prior is assumed to be nonatomic or discrete on a finite support. In addition, these two opposite cases are often stated and proved separately and in separate papers. For example, it is by now well understood that the extreme points of the feasible set are the bi-pooling distributions, but this result is only known for nonatomic priors (Arieli et al. (2023) also assuming full support, Kleiner et al. (2021) for any nonatomic prior). If the objective is concave-to-convex, the optimal information structure is a censorship policy, which is known both for discrete priors on finite support (Alonso and Câmara, 2016; Kolotilin et al., 2024) and for priors admitting a strictly positive density (Kolotilin et al., 2022), but not in the general case.

While such regularity assumptions are both natural and perfectly understandable in the development of a theory, they are also ultimately undesirable, both mathematically and for applications. From a mathematical point of view, the restrictions do not seem essential to the formulation of the problem or to much of the intuition behind its solution. A formal generalization is therefore natural, and helpful to clarify how far a perhaps seductive intuitive generalization actually carries. Furthermore, a general treatment brings together two common cases in applications: persuasion with a binary state, usually treated by concavification following Kamenica and Gentzkow (2011), and persuasion with a continuously distributed state and mean-dependent payoffs, usually treated with duality techniques. From a practical point of view, many applications are partially discrete. The information available to a designer may consist of coarse grades or ratings; it may come from an empirical distribution which is finitely supported; and it may feature mass points due to missing or top-coded data, for example in stress tests, credit scores or other settings involving other sensitive information. Even when the information of the designer is perfect, strategic bunching at regulatory thresholds can likewise create mass points in otherwise continuous distributions. Moreover, even in the most regular cases, optimal policies typically produce distributions outside the class of priors covered by the structural characterizations: nontrivial bi-pooling creates atoms at the pooled posterior means, and

a censorship policy creates an atom corresponding to the censored states. An intermediary who subsequently garbles this information therefore faces an atomic distribution even if the original prior was atomless. Models of sequential disclosure, information intermediaries dealing with designed information, or competing designers whose choices involve further garbling hence provide an additional motivation for a general theory. The aim of this paper is to provide such a general development.

We study the solutions of the problem

$$\sup_{G \in \mathrm{MPC}(F)} \int u \, \mathrm{d}G \tag{P}$$

for an upper semicontinuous u and a general $F \in \Delta([0,1])$, by studying the geometry of the set of mean-preserving contractions $\mathrm{MPC}(F)$. We first provide a canonical decomposition of elements of $\mathrm{MPC}(F)$ into pooling intervals and constrained regions (Theorem 1). Any $G \in \mathrm{MPC}(F)$ is a mean-preserving contraction of F when both are restricted to each pooling interval—where care needs to be taken to assign atoms at the boundary of these intervals—and coincides with F on the mass that is not assigned to any pooling interval. With this decomposition at hand, we can define our generalized notion of bi-pooling: A distribution $G \in \mathrm{MPC}(F)$ is bi-pooling if its restriction to any pooling interval is supported on at most two points (Definition 2). We also define a second notion, split rigidity. A distribution $G \in \mathrm{MPC}(F)$ is split-rigid if there is no other distribution in $\mathrm{MPC}(F)$ whose support is a subset of that of G and whose set of points where the mean-preserving contraction constraint is binding is a superset compared to that of G (Definition 3).

We then study the geometry of $\mathrm{MPC}(F)$. We show that every extreme point is bi-pooling, but in contrast with the case of atomless F, the converse is not generally true. The reason is that bi-pooling only restricts the behavior inside of pooling intervals, but not how atoms at the boundary of pooling intervals are split between them (Example 1). Not every extreme point needs to be exposed (Example 2); again, this contrasts with the case of atomless F where these notions coincide. We then show that an extreme point is exposed if and only if it is split-rigid (Theorem 2).

We then apply our results to the optimization problem (P). We show that there exists a bi-pooling solution (Theorem 3) and if u is continuous a unique solution needs to be split-rigid (Corollary 3). We discuss how working with atomic prior distributions allows us to consider the finite-state persuasion problem as a special case of moment persuasion. Every bi-pooling distribution can moreover be implemented by a simple information policy: moving to quantile space in order to split atoms, we construct both a monotone and a pure implementation (Proposition 5). Then, we study how shape restrictions on the objective simplify the problem; in particular, we show that generalized censorship policies are optimal when the objective is once continuously differentiable and S-shaped, without

any condition on the prior distribution, and that the optimum is unique when the S-shape is strict (Proposition 7).

1 Setup

Consider the space of finite nonnegative Borel measures on [0,1] and denote it by $\Delta([0,1])$. We will take all measures in this paper to belong to this set unless explicitly introduced as a signed measure. Throughout, we denote by $F \in \Delta([0,1])$ a fixed but arbitrary *probability* measure, also called the prior. [1] We use the same letter for a measure and its cumulative distribution function (CDF), writing

$$H(x) = H([0,x]), \qquad H(x-) = H([0,x)), \qquad \Delta H(x) = H(x) - H(x-)$$

for the mass of $[0,x]$, the mass of $[0,x)$, and the atom at x, respectively. We write $m_H = H([0,1])$ for the total mass and, when $m_H > 0$, we write

$$e_H = \frac{1}{m_H} \int_{[0,1]} x \, \mathrm{d}H(x)$$

for the mean. We use $\delta_x \in \Delta([0,1])$ for the unit measure on $x \in [0,1]$. We equip the space of finite signed Borel measures on $[0,1]$ with the weak* topology, which is induced by the integration against continuous functions. Unless otherwise specified, all topological statements about measures refer to this topology.

It will be useful to work with a measure H in terms of its *potential* defined as

$$R_H(x) = \int_{[0,1]} (x-y)_+ \, \mathrm{d}H(y), \qquad x \in [0,1].$$

The following properties of the potential are standard.

Lemma 1 (Potential calculus)**.** *Let H be a finite Borel measure on $[0,1]$. Then*

1. *$R_H(x) = \int_0^x H(t) \, \mathrm{d}t$ for all $x \in [0,1]$. In particular R_H is convex, nondecreasing, m_H-Lipschitz, and $R_H(0) = 0$.*
2. *$R_H(1) = m_H - \int x \, \mathrm{d}H(x)$.*
3. *R_H determines H on $[0,1)$; together with m_H it determines H. The same holds for finite signed measures.*

[1] The restriction to the unit interval (as opposed to any compact interval) is made to simplify notation and without loss of generality. We allow unnormalized measures because restrictions and components of a probability measure need not have unit mass. The normalization of the prior F simplifies notation without loss of generality, noting that if G is a mean-preserving contraction of F then $1/m G$ is a mean-preserving contraction of $1/m F$, for any $m > 0$.

The proof of the Lemma is given in the Appendix. We can now define a mean-preserving contraction.

Definition 1. Let μ, ν be finite Borel measures on $[0,1]$ with $m_\nu = m_\mu$. We say that ν is a **mean-preserving contraction** of μ, and write $\nu \in \mathrm{MPC}(\mu)$ if

$$R_\nu(x) \leq R_\mu(x) \quad \text{for all } x \in [0,1], \qquad \text{and} \qquad R_\nu(1) = R_\mu(1).$$

Note that the boundary condition implies that the two distributions indeed have the same first moment (by Lemma 1 (ii)), justifying the terminology. We will be interested in the geometry of the set $\mathrm{MPC}(F)$. This set is exactly the set of posterior-mean distributions of signals about F (Strassen, 1965). Furthermore, the order is equivalent to the standard convex order.

Lemma 2. *Let μ, ν be finite Borel measures on $[0,1]$ with $m_\nu = m_\mu$. Then $\nu \in \mathrm{MPC}(\mu)$ if and only if*

$$\int \varphi \,\mathrm{d}\nu \leq \int \varphi \,\mathrm{d}\mu \qquad \textit{for every continuous convex } \varphi : [0,1] \to \mathbb{R}.$$

Optimization We will be interested in the solutions of the optimization problem

$$\sup_{G \in \mathrm{MPC}(F)} \int u \,\mathrm{d}G \tag{P}$$

where $u : [0,1] \to \mathbb{R}$ denotes a fixed upper semicontinuous function and F denotes a fixed but arbitrary probability measure ("prior") on $[0,1]$. Note that the integral in (P) may be $-\infty$ for some measure. The supremum always takes a real value as it is bounded above by $\max_{x \in [0,1]} u(x)$ and since $\delta_{e_F} \in \mathrm{MPC}(F)$ is feasible and attains the value $u(e_F) \in \mathbb{R}$.

Lemma 3 (Topological structure)**.** $\mathrm{MPC}(F)$ *is a nonempty, convex, and weak* compact subset of* $\Delta([0,1])$ *for any probability measure* $F \in \Delta([0,1])$.

2 Bi-Pooling Distributions

A certain amount of bookkeeping is required to define bi-pooling in the general case and to work with the resulting decomposition. In this section, we develop the tools and notation required, and will close with the definition of (generalized) bi-pooling and split rigidity.

Fix a $G \in \mathrm{MPC}(F)$ and define the *potential gap* as the slack in the defining constraint of MPC,

$$D_G(x) := R_F(x) - R_G(x) = \int_0^x \Big(F(t) - G(t)\Big)\,\mathrm{d}t, \qquad x \in [0,1].$$

By Lemma 1, D_G is continuous (indeed 1-Lipschitz), nonnegative, and $D_G(0) = D_G(1) = 0$. The open set

$$U_G \coloneqq \{x \in (0,1) : D_G(x) > 0\}$$

is hence a countable union of pairwise disjoint maximal open intervals, which we call the *pooling intervals* of G (given F).[2] The complement

$$K_G = [0,1] \setminus U_G$$

is the *binding set*; it is closed and contains 0 and 1. The potential gap D_G has one-sided derivatives everywhere, which we denote as

$$D'_G(x+) = F(x) - G(x) \quad (x < 1), \qquad D'_G(x-) = F(x-) - G(x-) \quad (x > 0). \tag{1}$$

For a pooling interval $I = (a, b)$, define

$$\lambda_I^a \coloneqq D'_G(a+) \in [0, \Delta F(a)], \qquad \lambda_I^b \coloneqq -D'_G(b-) \in [0, \Delta F(b)]$$

The numbers λ_I^a and λ_I^b are nonnegative since the right (left) derivative of D_G is nonnegative (nonpositive) at the beginning (end) of a pooling interval. They denote the mass of the endpoint atoms of F that we assign to the interval I (as opposed to the complement K_G or another pooling interval adjacent to a or b). Concordantly, we define the *prior measure* and *posterior measure* on I as the measures

$$F_I \coloneqq F|_{(a,b)} + \lambda_I^a\, \delta_a + \lambda_I^b\, \delta_b, \qquad G_I \coloneqq G|_{(a,b)}. \tag{2}$$

Remark. Note that λ_I^a and λ_I^b are fully determined by G, through the one-sided slopes of its potential gap. Note also that the definition assigns atoms of G at the binding points a and b not to (any) G_I—they will be part of the residual—while any atom of G inside a pooling interval belongs to that interval's posterior component.

Theorem 1 (Canonical decomposition)**.** *Let $G \in \mathrm{MPC}(F)$ with pooling intervals $\{I\}$. Then:*

1. *$G_I \in \mathrm{MPC}(F_I)$ for every pooling interval I; in particular $m_{G_I} = m_{F_I}$ and $e_{G_I} = e_{F_I}$, and the common mass is always positive.*

2. *The prior and posterior residuals coincide:*

$$F - \sum_I F_I = G - \sum_I G_I = G|_{K_G} \geq 0.$$

[2] This terminology anticipates the definition of bi-pooling below, which will require that on every pooling interval mass is indeed pooled to at most two support points. Formally, the intervals are simply the maximal connected components of U_G.

Figure 1 illustrates the decomposition for the distribution $G_{1/2}$ of Example 1 below. There is an atom at the binding point $x = \frac{1}{2} \in K_{G_{1/2}}$ which is split equally between both prior components; each posterior component is the restriction of $G_{1/2}$ to its pooling interval. There is no residual mass.

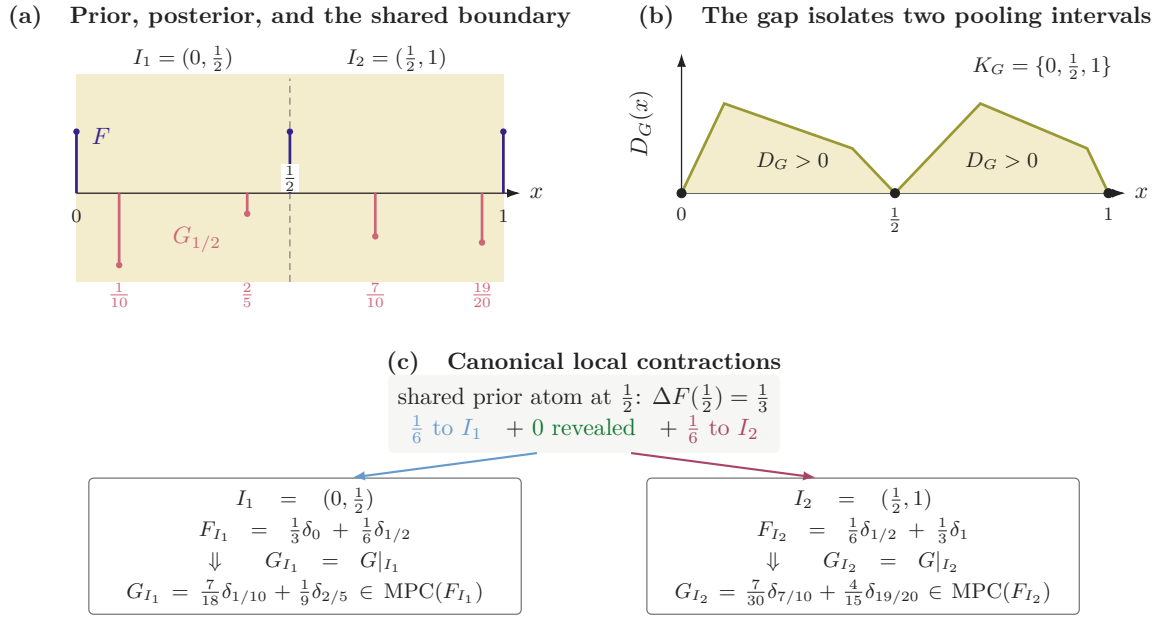

Figure 1: A canonical decomposition with a shared boundary atom. Panel (a) shows the prior above the axis and the posterior below it; panel (b) identifies the two pooling intervals through the potential gap; panel (c) displays the two local mean-preserving contractions. The atom of mass 1/3 at 1/2 is divided equally between the two canonical prior components.

Corollary 1 (Atom budget)**.** *For every interior binding point $x \in K_G \cap (0,1)$, let $I_{x-} = (a, x)$ denote the pooling interval that has x as its right endpoint (if it exists), and let $I_{x+} = (x, b)$ denote the pooling interval that has x as its left endpoint (if it exists). Then*

$$\Delta F(x) = \lambda^x_{I_{x-}} + \Delta G(x) + \lambda^x_{I_{x+}}.$$

In other words, the atom at a binding point x is exactly split between the preceding pooling interval, the atom of G at x and the succeeding pooling interval. If one or both of these intervals do not exist, the associated λ are taken as zero: no mass can be shifted to the left/right if there is no immediately preceding/succeeding interval.

Definition 2 (Bi-pooling)**.** A distribution $G \in \mathrm{MPC}(F)$ is ***bi-pooling*** if

$$|\operatorname{supp}(G_I)| \le 2 \qquad \text{for every pooling interval } I \text{ of } G.$$

Throughout, we will refer to this generalized notion simply as bi-pooling; we retain the qualifier "generalized" only when comparison with the nonatomic definition in the literature requires clarification.

A distribution G is bi-pooling exactly when, on each pooling interval, it consists of one or two atoms, whose masses and locations satisfy the mass and barycenter constraints of Theorem 1(i).

Remark 1 (Nonatomic case). If F is nonatomic, we have $\lambda_I^a = \lambda_I^b = 0$, so $F_I = F|_{(a,b)}$ and the residual identity reads $F|_{K_G} = G|_{K_G}$: A generalized bi-pooling G reveals the state outside the pooling intervals, and on each pooling interval the prior mass is contracted to at most two posterior means. This is precisely the bi-pooling notion of Arieli et al. (2023) and Kleiner et al. (2021).

We will see that the presence of mass points gives bi-pooling distributions more flexibility than in the atomless case. In particular, there can be two distinct bi-pooling distributions with nested support that have the same set of points where the MPC constraint binds. The following more restrictive definition rules out this case.

Definition 3 (Split rigidity)**.** A distribution $G \in \mathrm{MPC}(F)$ is *split-rigid* if the only $H \in \mathrm{MPC}(F)$ with

$$\mathrm{supp}(H) \subseteq \mathrm{supp}(G) \qquad \text{and} \qquad R_H(x) = R_G(x) = R_F(x) \ \text{ for all } x \in K_G$$

is G itself.

3 Exposed and Extreme Points

In this section, we derive our results about the geometry of $\mathrm{MPC}(F)$ for an arbitrary prior F.

Recall that a point x in a convex set X is called extreme if there are no $x' \neq x'' \in X$ and $\alpha \in (0,1)$ such that $x = \alpha x' + (1-\alpha)x''$. We will denote the set of extreme points as $\mathrm{ext}(X)$. A point x in a convex set X is called exposed if there exists a continuous linear functional y such that $y(x) > y(x')$ for all other $x' \in X$. In our setting, recall that the continuous linear functionals are the maps $H \mapsto \int u \,\mathrm{d}H$ with $u \in C[0,1]$. Clearly, every exposed point is extreme but the converse need not be true.

We are now ready to state our main theorem.

Theorem 2. *For every probability measure* $F \in \Delta([0,1])$

$$\{\textit{exposed}\} = \{\textit{split-rigid}\} \subseteq \{\textit{extreme}\} \subseteq \{\textit{bi-pooling}\} \tag{3}$$

as subsets of $\mathrm{MPC}(F)$*. Both inclusions are strict for some atomic priors.*

Proof. Deferred to Sections 3.1-3.5. □

The theorem provides a sharp characterization of the exposed points of MPC(F) in terms of split rigidity. The classic and more interpretable bi-pooling property is only a necessary condition for extremality. Interestingly, not only are there bi-pooling distributions that are not extreme, there are also extreme points that are not exposed. Neither phenomenon can arise under the assumption of the prior literature that the prior F admits a continuous cdf.

We can recover this equivalence under a local condition. All complications creating strict inclusions between these notions are caused by atoms of the prior distribution at the boundary of the pooling intervals.

Proposition 1 (No atoms on the pooling boundary)**.** *Let $G \in \mathrm{MPC}(F)$ be bi-pooling. If*

$$\Delta F(x) = 0 \qquad \text{for every } x \in \partial K_G,$$

where ∂K_G is the relative boundary of K_G in $[0,1]$, then G is split-rigid, and hence exposed and extreme.

Proof. Fix a bi-pooling distribution $G \in \mathrm{MPC}(F)$ and let $H \in \mathrm{MPC}(F)$ satisfy the two conditions in Definition 3. We need to show that $H = G$.

First, note that at any $x \in \partial K_G \subset K_G$, we have $R_H(x) = R_G(x) = R_F(x)$, so Lemma 5 (from the appendix), using $\Delta F(x) = 0$ gives

$$H(x-) = H(x) = G(x-) = G(x) = F(x-) = F(x). \tag{4}$$

In the interior of K_G, the equality $R_H = R_F = R_G$ implies that the cdfs coincide by absolute continuity. Hence, the cdfs of H and G agree on K_G.

Fix a pooling interval $I = (a, b)$. Since $\mathrm{supp}(H) \subset \mathrm{supp}(G)$ and G is bi-pooling both $H|_I$ and G_I are supported on the same set of at most two points. Their masses agree since $a, b \in \partial K_G$ and hence by (4) we know that $H(b-) - H(a) = G(b-) - G(a)$, and their first moments agree because $R_H(b) - R_H(a) = R_G(b) - R_G(a)$. A measure on at most two fixed points is uniquely determined by its mass and first moment, so $H|_I = G_I$. This holds for every pooling interval. Hence, the measures agree also on U_G, and $H = G$. □

Corollary 2 (Nonatomic priors)**.** *Let F be nonatomic. Then every bi-pooling $G \in \mathrm{MPC}(F)$ is split-rigid. Consequently the four classes of Theorem 2 coincide:*

$$\{\textit{split-rigid}\} = \{\textit{exposed}\} = \{\textit{extreme}\} = \{\textit{bi-pooling}\}.$$

Corollary 2 recovers the extreme-point and exposed-point characterizations of Arieli et al. (2023) and Kleiner et al. (2021). Nonatomicity is precisely the assumption of Theorem 2 in Kleiner et al. (2021), which requires a continuous CDF.

We will now proceed to prove each statement in the inclusion chain (3) in a separate subsection, proceeding right to left. The reader less interested in the mathematical details of the constructions may skip directly to Section 4 where we apply the result to optimization problems commonly arising in information design.

3.1 Every Extreme Point is Bi-Pooling

Proposition 2. *Every extreme point of* $\mathrm{MPC}(F)$ *is bi-pooling.*

The proof adapts the three-packet perturbation of Arieli et al. (2023, Lemma 5) to arbitrary priors and the generalized definitions of the present paper. In CDF coordinates, the perturbation is closely related to the local level-shift arguments in the proof of Theorem 2 of Kleiner et al. (2021).

Proof. Let $G \in \mathrm{MPC}(F)$ and suppose some pooling interval $I = (a, b)$ has $|\operatorname{supp}(G_I)| \geq 3$; we will produce a nontrivial decomposition $G = \frac{1}{2}(G + \theta\eta) + \frac{1}{2}(G - \theta\eta)$ within $\mathrm{MPC}(F)$.

First, let us construct the perturbation η. Pick $x_1 < x_2 < x_3$ in $\operatorname{supp}(G_I) \cap (a, b)$; since G_I lives on an open interval, we may take the x_j to be points such that every neighborhood of them has positive measure under G_I. We can then pick a $\delta > 0$ sufficiently small and set $B_j = (x_j - \delta, x_j + \delta)$ such that the B_j are pairwise disjoint, have closures inside (a, b), are ordered ($\sup B_1 < \inf B_2$, $\sup B_2 < \inf B_3$), and each has $G(B_j) > 0$. Let G_j be the normalized restriction of G to B_j and $m_j = \int x \, \mathrm{d}G_j(x) \in \overline{B_j}$ the mean; the ordering of the sets gives $m_1 < m_2 < m_3$.

Choose $\beta \in (0, 1)$ with $\beta m_1 + (1 - \beta)m_3 = m_2$ and set

$$\eta = \beta G_1 + (1 - \beta)G_3 - G_2.$$

Then η is a nonzero signed measure with $\eta([0, 1]) = 0$ and $\int x \, \mathrm{d}\eta = 0$. For $\theta \leq \min\Big(G(B_2),\, G(B_1)/\beta,\, G(B_3)/(1 - \beta)\Big)$ both $G + \theta\eta$ and $G - \theta\eta$ are nonnegative measures of total mass one: the negative parts θG_2, respectively $\theta\beta G_1 + \theta(1 - \beta)G_3$, are dominated by the restrictions of G to the corresponding sets.

We now show that the perturbation is feasible. Consider $R_\eta(x)$, the potential of the signed measure η. For $x \leq \inf B_1$, $\eta([0, x]) = 0$, so $R_\eta(x) = 0$. For $x \geq \sup B_3$, we have $\eta([0, x]) = 0$, so R_η is constant there, with value $R_\eta(1) = \eta([0, 1]) - \int x \, \mathrm{d}\eta = 0$ by Lemma 1(ii). Hence $R_\eta(x)$ is supported on the compact set $J = [\inf B_1, \sup B_3] \subset I$, and $C = \max_{x \in J} |R_\eta(x)| < \infty$. Since D_G is continuous and strictly positive on J, $\epsilon_0 = \min_J D_G > 0$. For $\theta \leq \epsilon_0 / C$ (any θ if $C = 0$), we have

$$R_{G \pm \theta\eta}(x) = R_G(x) \pm \theta R_\eta(x) \leq R_G(x) + \theta C \cdot \mathbf{1}_J(x) \leq R_F(x) \qquad \text{for all } x,$$

and $R_{G\pm\theta\eta}(1) = R_G(1) = R_F(1)$. So both perturbed measures lie in MPC(F), they are distinct, and they average to G. Hence G is not extreme, which proves the contrapositive of the claim. □

3.2 Not Every Bi-Pooling Distribution is an Extreme Point

For nonatomic priors, bi-pooling *characterizes* extremality (Corollary 2 above). With atoms, the inclusion is strict. The reason is that bi-pooling restricts the behavior only inside the pooling intervals, it does not restrict how an atom on the binding set is split between its two neighboring pooling intervals.

This becomes visually apparent in the language of potentials (see Figure 2). The binding set consists of the contact points where R_G touches R_F from below. If F is nonatomic, R_F is differentiable, and contact forces both one-sided slopes of R_G to equal its slope. For nonatomic F, therefore, the contact condition exactly pins down the behavior of R_G on the binding set; and bi-pooling pins down the behavior on the pooling intervals. When F has an atom on the binding set, however, R_F has a corner, and contact does not determine the slopes of R_G uniquely. Different slopes correspond to different allocations of the prior atom between revelation and pooling into the adjacent pooling intervals. Bi-pooling restricts the number of posterior locations inside each pooling interval, but does not by itself rule out these changes in allocation. They can be used to construct examples of bi-pooling distributions that can be symmetrically perturbed within the set of mean-preserving contractions, and which are hence not extreme.

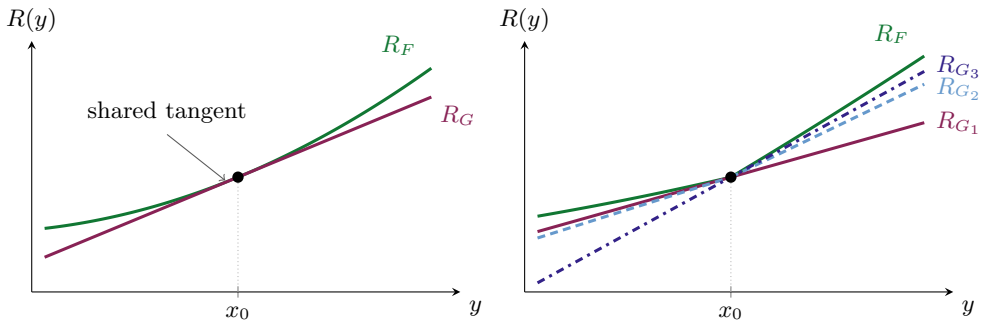

Figure 2: Contact condition (left) and degrees of freedom at an atom (right).

The following example demonstrates this phenomenon.

Example 1 (A split atom creates non-extreme bi-pooling distributions)**.** Let

$$F = \tfrac{1}{3}\delta_0 + \tfrac{1}{3}\delta_{1/2} + \tfrac{1}{3}\delta_1.$$

For $t \in [1/4, 8/9]$ define

$$G_t = \frac{4-t}{9}\,\delta_{1/10} + \frac{4t-1}{9}\,\delta_{2/5} + \frac{8-9t}{15}\,\delta_{7/10} + \frac{2+4t}{15}\,\delta_{19/20}.$$

Then:

1. $G_t \in \mathrm{MPC}(F)$, with binding set $K_{G_t} = \{0, 1/2, 1\}$ and pooling intervals $(0, 1/2)$ and $(1/2, 1)$;

2. every G_t is bi-pooling, with the same support $\{1/10, 2/5, 7/10, 19/20\}$ for every $t \in (1/4, 8/9)$;

3. the weights are affine in t, so for every $t \in (1/4, 8/9)$ and small $\varepsilon > 0$,

$$G_t = \tfrac{1}{2}G_{t-\varepsilon} + \tfrac{1}{2}G_{t+\varepsilon},$$

and no G_t is an extreme point of $\mathrm{MPC}(F)$ for $t \in (1/4, 8/9)$.

Here, $\lambda^{1/2}_{(0,1/2)} = t/3$ and $\lambda^{1/2}_{(1/2,1)} = (1-t)/3$ are the shares of the F-atom at $1/2$ assigned to the left and right pooling intervals. Intuitively, within each pooling interval, the assigned mass is accommodated on the two atoms and shifted between them to ensure that $G_I \in \mathrm{MPC}(F_I)$ in accordance with (2) and Theorem 1. Figure 3 illustrates the distributions (left) and the degree of freedom in potential space (right). Detailed calculations are in Appendix B.1.

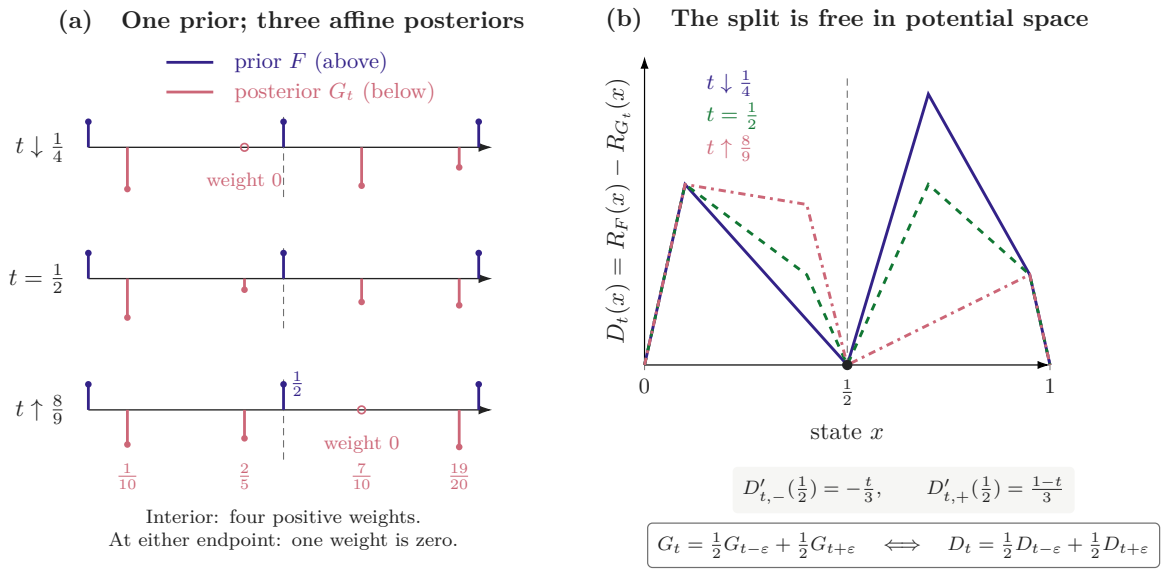

Figure 3: The family G_t of Example 1. Panel (a) shows three members of the family, including its two limiting cases. Panel (b) plots the associated potentials.

Remark 2 (Atomless Approximation). One possible approach to deal with atomic priors is by approximation. For any atomic F, there is a parameterized family F_ϵ such that $F_\epsilon \to F$ weakly (indeed, many such families). It might therefore seem natural to consider the limits of extreme points of $\mathrm{MPC}(F_\epsilon)$. Example 1 also demonstrates that this approach cannot be taken lightly as extremality is not preserved by the limit. In Appendix B.1, we construct such a nonatomic F_ϵ for the prior above, where $G_{1/2} \in \mathrm{extMPC}(F_\epsilon)$ for sufficiently small $\epsilon > 0$, yet, by the above, $G_{1/2} \notin \mathrm{extMPC}(F)$.

3.3 Not Every Extreme Point is Split-Rigid

We demonstrate this strict inclusion by example. As is familiar from the previous example, bi-pooling does not restrict the movement of mass from atoms of F on the binding set into adjacent pooling intervals. The example exploits this fact to construct a $G \in \mathrm{ext}(\mathrm{MPC}(F))$ that is not split-rigid and hence not exposed.

Let us first give an intuitive explanation. Consider an F with a segment with positive density adjacent to a mass point, and a mean-preserving contraction that pools the segment of positive density but keeps the mass point intact. Such a mean-preserving contraction (at least to the degree we described it locally) is bi-pooling and may be extreme. But how could it be supported? Pooling the density region is optimal only if the states just below the atom are worth weakly more inside the pool than revealed. Keeping the atom intact is optimal only if the atom is worth weakly more revealed than inside the pool. With a continuous integrand, the two comparisons coincide in the limit, so the atom must be exactly indifferent: moving some of its mass into the pool is then optimal as well. Such a distribution can therefore be optimal, but never uniquely so. This idea gives an extreme point that cannot be exposed by any continuous u. Note that it could be exposed by an upper semicontinuous u by placing an upward discontinuity at the atom. (Indeed, this is the flavor of Example 5 in the Appendix.) The construction below is more careful and more powerful and rules this out by placing an additional mass point: the extreme point is not exposed by any linear functional.

Example 2 (One-way gate)**.** Let

$$F = \frac{1}{2}\mathsf{U}[0, \frac{1}{2}] + \frac{1}{4}\delta_{1/2} + \frac{1}{4}\delta_1,$$

where $\mathsf{U}[0, \frac{1}{2}]$ denotes the uniform distribution on $[0, \frac{1}{2}]$, and

$$G = \frac{1}{4}\delta_{1/6} + \frac{1}{4}\delta_{1/3} + \frac{1}{4}\delta_{1/2} + \frac{1}{4}\delta_1.$$

Then $G \in \mathrm{MPC}(F)$ is bi-pooling and extreme, but it is not split-rigid. Furthermore, it is not exposed. Indeed, it is not exposed in any topology as no real-valued linear functional exposes G.

The distribution G pools the uniform component of F on the two points $1/6$ and $1/3$, while revealing the atoms at $1/2$ and 1. That G is bi-pooling is therefore immediate. That it is extreme requires a more careful argument (provided in Appendix B.2). Intuitively, it is the endpoint of a nontrivial segment of $\mathrm{MPC}(F)$ with the same support and the same binding set. To see this more formally, consider the balanced signed measure

$$\xi = -\delta_{1/6} + 2\delta_{1/3} - \delta_{1/2}.$$

Appendix B.2 shows that

$$\{H \in \mathrm{MPC}(F) : \mathrm{supp}(H) \subseteq \mathrm{supp}(G),\ R_H = R_F \text{ on } K_G\} = \left\{G + t\xi : 0 \le t \le \frac{1}{4}\right\}. \quad (5)$$

Thus G is not split-rigid. Geometrically, however, G is the endpoint of this segment: the movement in the direction $+\xi$ is feasible, whereas the movement in the direction $-\xi$ immediately violates the potential constraint just to the left of $1/2$. Therefore, G cannot be written as a convex combination in $\mathrm{MPC}(F)$.

Despite that, G is not exposed. Even though the straight movement in direction $-\xi$ is infeasible, there is a feasible curved path having $-\xi$ as its first-order direction. Namely, let

$$\rho = -3\delta_{1/6} + 3\delta_{1/3} + \delta_{1/2} - \delta_1$$

and

$$J_\varepsilon = G - \varepsilon\xi + \varepsilon^2\rho.$$

For every sufficiently small $\varepsilon > 0$, $J_\varepsilon \in \mathrm{MPC}(F)$. If a linear functional Λ is maximized at G, feasibility of $G + t\xi$ implies $\Lambda(\xi) \le 0$, while feasibility of J_ε implies that $\Lambda(G) - \Lambda(J_\varepsilon) = \varepsilon\Lambda(\xi) - \varepsilon^2\Lambda(\rho) \ge 0$ for all small $\varepsilon > 0$ which implies that $\Lambda(\xi) \ge 0$. Thus $\Lambda(\xi) = 0$, and Λ is constant on the nontrivial segment $G + t\xi$. No linear functional can therefore expose G.

We provide a distinct example of a non-exposed extreme point as Example 4 in Appendix B.3. The example is less tangible as it relies on an accumulation of point masses, but it gives a qualitatively distinct reason why an extreme point is not exposed.

We also show that the two reasons an extreme point fails to be exposed are in a sense exhaustive. Proposition 8 in Appendix B.4 shows that if all atoms of F are isolated (ruling out Example 2) and there are only finitely many pooling intervals (ruling out Example 4), all extreme points are split-rigid and exposed. In particular, the three classes coincide whenever F has finite support, since a prior with n atoms allows at most $n - 1$ pooling

intervals (Corollary 5). Example 1, which satisfies both these conditions, shows that they are not sufficient to ensure that all bi-pooling distributions are extreme.

Remark 3. Note that the argument showing that G is not exposed in Example 2 uses only linearity and did not rely on continuity. It hence holds for any topology on $\Delta([0,1])$. The argument for Example 4 does pass to a limit in total variation, so it rules out every linear functional that is continuous in total variation, though not every algebraic linear functional. Indeed, Appendix B.3 exhibits a linear functional that exposes the distribution of Example 4, so it is exposed under the finest locally convex topology. The failure of exposedness is thus algebraic in Example 2 and topological in Example 4.

3.4 Every Split-Rigid Distribution is Exposed

Proposition 3 (Split rigidity implies exposedness)**.** *Every split-rigid $G \in \mathrm{MPC}(F)$ is an exposed point of* $\mathrm{MPC}(F)$.

Proof. To establish this result, we will construct a continuous (indeed, Lipschitz) $u : [0,1] \to \mathbb{R}$ such that G is the unique maximizer of $H \mapsto \int u \,\mathrm{d}H$ over $\mathrm{MPC}(F)$, for a given split-rigid G.

Recall that the binding set K_G is a nonempty closed subset of $[0,1]$. Choose a sequence $(t_n)_{n\in\mathbb{N}}$ in K_G whose range is dense in K_G and let $\Gamma = \sum_n 2^{-n}\delta_{t_n}$, a finite positive measure with $\operatorname{supp}\Gamma = K_G$. Recall that, by Lemma 1, $R_\Gamma(x) = \int_{[0,1]}(x-t)_+ \,\mathrm{d}\Gamma(t)$ is convex and Lipschitz continuous.

For any $H \in \mathrm{MPC}(F)$, Fubini–Tonelli and the decomposition $(x-t)_+ = (x-t)+(t-x)_+$ imply that

$$\begin{aligned}\int R_\Gamma(x)\,\mathrm{d}H(x) &= \int \left(\int (x-t)_+ \,\mathrm{d}\Gamma(t)\right) \mathrm{d}H(x)\\ &= \int \left(\int (x-t) + (t-x)_+ \,\mathrm{d}H(x)\right) \mathrm{d}\Gamma(t)\\ &= \int e_H - t + R_H(t)\,\mathrm{d}\Gamma(t)\end{aligned}$$

Since every feasible distribution is a probability measure with mean e_F, we have for all $H \in \mathrm{MPC}(F)$

$$\int R_\Gamma(x)\,\mathrm{d}F(x) - \int R_\Gamma(x)\,\mathrm{d}H(x) = \int_{K_G} \Big(R_F(t) - R_H(t)\Big)\,\mathrm{d}\Gamma(t) \geq 0, \tag{6}$$

since the integrand is nonnegative and the integral can be taken over K_G because $\Gamma(t)$ is supported only on K_G. Applying (6) to $H = G$, we see that it holds with equality, since R_G agrees with R_F on K_G by definition of the binding set.

We are now ready to define the exposing functional. Let $d(\cdot, \operatorname{supp} G) : [0,1] \to [0,1]$ denote the Euclidean distance from the support of G. Note that it is continuous (indeed,

1-Lipschitz). Then, let

$$u(x) = R_\Gamma(x) - d(x, \operatorname{supp} G).$$

Note that u is continuous, $u \le R_\Gamma$ everywhere, and $u = R_\Gamma(x)$ exactly on $\operatorname{supp} G$. Since $G([0,1] \setminus \operatorname{supp} G) = 0$, we have that

$$\int u \, \mathrm{d}G = \int R_\Gamma(x) \, \mathrm{d}G = \int R_\Gamma(x) \, \mathrm{d}F.$$

For an arbitrary $H \in \mathrm{MPC}(F)$, combining $u \le R_\Gamma$ with (6), we get that

$$\int u \, \mathrm{d}H \le \int R_\Gamma \, \mathrm{d}H \le \int R_\Gamma \, \mathrm{d}F = \int u \, \mathrm{d}G, \tag{7}$$

so G is a maximizer. It remains to establish uniqueness.

Let $H \in \mathrm{MPC}(F)$ attain the maximum. Then, both inequalities in (7) must hold with equality. Equality in the first gives $\int d(x, \operatorname{supp} G) \, \mathrm{d}H(x) = 0$, so H is concentrated on the closed set $\operatorname{supp}(G)$, i.e. $\operatorname{supp}(H) \subseteq \operatorname{supp}(G)$. Equality in the second, by (6), gives

$$\int_{K_G} \big(R_F(t) - R_H(t)\big) \, \mathrm{d}\Gamma(t) = 0.$$

The integrand is continuous, nonnegative, and integrates to zero against a measure whose support is all of K_G; hence it vanishes on K_G. Therefore, we have $R_H = R_F$ on K_G. Split rigidity therefore implies that $H = G$. □

3.5 Every Exposed Distribution is Split-Rigid

It remains to show that every exposed distribution is split-rigid. This argument relies on duality techniques, which we relegate to the appendix. In the main text, we give the key substantive lemma which states that shrinking an optimizer's support and preserving its binding constraints preserves optimality.

Lemma 4 (Preservation of optimality)**.** *Let $u \in C[0,1]$, and suppose that $G \in \mathrm{MPC}(F)$ maximizes $\int u \, \mathrm{d}J$ over $J \in \mathrm{MPC}(F)$. If $H \in \mathrm{MPC}(F)$ satisfies*

$$\operatorname{supp}(H) \subseteq \operatorname{supp}(G), \qquad R_H = R_F \quad \text{on } K_G,$$

then H also maximizes $\int u \, \mathrm{d}J$ over $J \in \mathrm{MPC}(F)$.

The proof of the lemma is relegated to the Appendix. It approximates the continuous objective by Lipschitz functions and applies the supporting-price theorem of Dworczak and Martini (2019). After passing to the limit to recover a dual certificate and a complementary slackness formula for the original problem with a continuous but not necessarily Lipschitz objective, it shows that an H satisfying the conditions of the Lemma remains optimal.

Proposition 4. *Every exposed point of* $\mathrm{MPC}(F)$ *is split-rigid.*

Proof. Let G be exposed by $u \in C[0,1]$, and let $H \in \mathrm{MPC}(F)$ satisfy

$$\mathrm{supp}(H) \subseteq \mathrm{supp}(G), \qquad R_H = R_F \quad \text{on } K_G.$$

By Lemma 4, H also maximizes the exposing objective. Hence $H = G$, and G is split-rigid. □

Together with Proposition 3, this establishes the equivalence of exposedness and split rigidity. Since the construction in that proposition gives a Lipschitz exposing objective, every exposed point can in fact be exposed by a Lipschitz function.

4 Optimization

The preceding section describes the geometry of $\mathrm{MPC}(F)$. We now use that geometry to study the solutions to optimization problems. We first show that an optimum is attained by a bi-pooling distribution for every upper semicontinuous objective and every prior, including atomic, mixed, and empirical priors. Second, we discuss how working with atomic prior distributions allows us to consider the finite-state persuasion problem as a special case of moment persuasion. Third, we show that every bi-pooling distribution can be implemented by two intuitive signal structures: One that is stochastically monotone and one that is deterministic as a function of the state when viewed in quantile space to split atoms. Then, we show that shape restrictions on the objective reduce the problem to a finite dimensional optimization over interval endpoints, atom shares, and posterior means. Finally, we show that generalized censorship policies are optimal when the objective is S-shaped.

4.1 Bi-Pooling and Split-Rigid Optima

Theorem 3 (Existence of optimal bi-pooling)**.** *Let* $u : [0,1] \to \mathbb{R}$ *be upper semicontinuous. Then*

$$\sup_{G \in \mathrm{MPC}(F)} \int u \, \mathrm{d}G \tag{P}$$

is attained, and it is attained by an extreme point of $\mathrm{MPC}(F)$*; in particular, by Proposition 2, by a bi-pooling distribution.*

Proof. An upper semicontinuous function on a compact metric space is bounded above and is the pointwise limit of a decreasing sequence of continuous functions $u_n \downarrow u$. For fixed G, monotone convergence gives $\int u \, \mathrm{d}G = \inf_n \int u_n \, \mathrm{d}G$, so $G \mapsto \int u \, \mathrm{d}G$ is an infimum of continuous functionals, hence upper semicontinuous with values in $[-\infty, \infty)$. It is

not identically $-\infty$, since the point mass at the prior mean $\delta_{e_F} \in \mathrm{MPC}(F)$ gives the finite value $u(e_F)$. By Lemma 3 and Weierstrass the maximum is attained, and the set of maximizers M is a nonempty, weak* compact, convex subset of $\mathrm{MPC}(F)$. It is also a face of $\mathrm{MPC}(F)$, because the objective is affine. By Krein–Milman, M has an extreme point, which is then extreme in $\mathrm{MPC}(F)$. The conclusion follows from Proposition 2. □

This theorem extends the insights of Arieli et al. (2023); Kleiner et al. (2021) to general prior distributions. When solving a linear optimization problem over $\mathrm{MPC}(F)$, we are guaranteed that there exists a bi-pooling solution.

Remark 4. The theorem also aims for generality on the objective, requiring only upper semicontinuity. Without a regularity condition of this kind, existence can fail. To see this, let $u = \mathbf{1}_{[0,1/2)}$ and $F = \frac{1}{2}\delta_0 + \frac{1}{2}\delta_1$. Then, $H_\epsilon = \frac{1}{1+2\epsilon}\delta_{1/2-\epsilon} + \frac{2\epsilon}{1+2\epsilon}\delta_1$, $\epsilon \in (0, \frac{1}{2})$ is in $\mathrm{MPC}(F)$ and $\int u \,\mathrm{d}H_\epsilon = \frac{1}{1+2\epsilon} \to 1$, yet the supremum is not attained, as no mean $1/2$ distribution can have all its mass on $[0, 1/2)$.

The theorem does not guarantee that all solutions are bi-pooling—just as in the smooth case, the objective may have flat segments that allow for many solutions even outside of extreme points—and it does not guarantee uniqueness. When the objective is continuous and the optimum is unique, however, we can further restrict the candidates to split-rigid distributions.

Corollary 3 (Unique Optima)**.** *If u is continuous and (P) has a unique maximizer G, then G is exposed and hence split-rigid.*

Proof. The linear functional $H \mapsto \int u \,\mathrm{d}H$ is continuous and uniquely maximized at G, so it exposes G. Hence, by Proposition 4, G is split-rigid. □

Remark 5 (Unique solutions under upper semicontinuous objectives). The continuity of u in Corollary 3 is essential. A distribution that is uniquely optimal for an upper semicontinuous objective is extreme, but need not be split-rigid or exposed. In particular, upward payoff jumps at atoms on the boundary of the prior's support hull can uniquely reward exact revelation of those atoms and thus uniquely select a distribution that fails split rigidity. Appendix B.5 gives an explicit bounded upper semicontinuous example with a unique maximizer that fails to be split-rigid.
In light of Remark 3, this means that among the non-exposed extreme points of $\mathrm{MPC}(F)$ there can both be distributions that cannot be exposed by any linear functional and distributions that can be exposed by refining the topology on $\Delta([0,1])$ to allow for bounded upper semicontinuous test functions.

Clearly, the extreme point logic also generalizes to convex objective functions. Examples of convex objective functions $\Phi : \Delta([0,1]) \to [-\infty, \infty)$ include upper envelopes of linear functionals, as in

$$\Phi(G) = \max_{i \le n} \int u_i \,\mathrm{d}G$$

for upper semicontinuous u_i, and general continuous convex aggregators, as in

$$\Phi(G) = \psi\Big(\int u_1 \,\mathrm{d}G, \ldots, \int u_n \,\mathrm{d}G\Big)$$

for continuous u_i and continuous convex ψ.

Corollary 4. *Let $\Phi : \Delta([0,1]) \to [-\infty, \infty)$ be convex and upper semicontinuous, and suppose that it is not identically $-\infty$ on $\mathrm{MPC}(F)$. Then Φ attains its maximum over $\mathrm{MPC}(F)$ at an extreme point, hence at a bi-pooling distribution.*

Proof. The maximum is finite and attained by upper semicontinuity and compactness. Choose N so large that the maximum exceeds $-N$, and apply Bauer's maximum principle to the real-valued convex upper semicontinuous function $\max\{\Phi, -N\}$. An extreme maximizer of the truncation also maximizes Φ. Proposition 2 completes the proof. □

4.2 Finite-State Persuasion

In the literature, there is sometimes an apparent disconnect between the tools and language used in persuasion problems with finite states ("concavification techniques") and with mean-dependent objectives. Being able to handle atomic prior distributions in the mean-dependent case allows for a (partial) unification of these approaches.

Our formulation covers finite-state persuasion whenever payoffs depend on a posterior belief only through its posterior mean. Let $\Theta = \{\theta_0 < \cdots < \theta_{n-1}\} \subset [0,1]$ denote the possible values of the scalar statistic whose posterior expectation determines the sender's payoff.[3] Let the prior assign probability $p_i > 0$ to θ_i, and write $F = \sum_i p_i \delta_{\theta_i}$. A payoff on posterior beliefs $q \in \Delta(\Theta)$ is *mean-measurable* if it has the form

$$w(q) = u(\sum_i q_i \theta_i).$$

The set of feasible distributions over posterior means is exactly $\mathrm{MPC}(F)$ (Strassen, 1965). Thus, finite-state persuasion with a mean-measurable payoff is exactly problem (P).

For the binary prior $F_0 = (1-p_1)\delta_0 + p_1\delta_1$, the posterior mean is simply the posterior probability of state $\theta_1 = 1$. Thus, mean-measurability is without loss of generality. Moreover, every probability measure G with mean p_1 belongs to $\mathrm{MPC}(F_0)$, feasibility reduces to the Bayes-plausibility condition. Within this feasible set, the extreme, exposed, and split-rigid distributions are exactly those supported on at most two points. Ordinary two-state concavification is therefore a special case of the present formulation. Beyond the conceptual unification, this perspective can help unify proofs of smooth and binary cases in applications (see e.g. Ostrizek, 2024).

[3] We have affinely rescaled this statistic to $[0,1]$, without loss of further generality.

For $n > 2$, an arbitrary payoff on posterior beliefs in general cannot be expressed as a function of a single posterior mean. It can, however, be expressed as a function of $n-1$ posterior moments: the posterior probabilities of the first $n-1$ states. Again, the prior distribution over these moments has finite support. General finite-state persuasion has an exact representation as an atomic multidimensional moment persuasion problem. Extending our analysis of the general prior case to this multidimensional problem is left for future research (see Dworczak and Kolotilin (2024) and Kleiner et al. (forthcoming) for approaches assuming more regularity for many of their results).

4.3 Implementation: From Bi-Pooling Distributions to Signals

We now show how to implement a (generalized) bi-pooling distribution by a simple information policy. The construction follows the bi-pooling implementations in Arieli et al. (2023); in order to generalize their approach we move into quantile space to be able to split the atoms of F. This requires us to first introduce the associated notation.

Quantile Notation If μ is a nonzero finite measure of mass m_μ, we write

$$Q_\mu(r) := \inf\{x : \mu([0,x]) \geq r\}, \qquad 0 < r < m_\mu,$$

for its quantile function. If r is uniform on $(0, m_\mu)$, then $Q_\mu(r)$ has distribution μ/m_μ, and $\int_0^{m_\mu} Q_\mu(r)\,\mathrm{d}r = m_\mu e_\mu$. We use the following quantile form of convex order: if $\mu \in \mathrm{MPC}(\nu)$ both measures of mass m, then, for every $s \in [0,m]$,

$$\int_0^s Q_\mu(r)\,\mathrm{d}r \geq \int_0^s Q_\nu(r)\,\mathrm{d}r, \qquad \int_s^m Q_\mu(r)\,\mathrm{d}r \leq \int_s^m Q_\nu(r)\,\mathrm{d}r. \tag{8}$$

From Distributions to Signals Take q uniform on $(0,1)$, so that the prior state is $Q_F(q)$. A *quantile signal* assigns to each q a lottery over posterior means $s \in [0,1]$. It implements G if s has distribution G and

$$\mathbb{E}[Q_F(q)|s] = s.$$

It is *pure* if s is deterministic given q, and *monotone* if the lottery over s is increasing in q in first-order stochastic dominance. If F has an atom at x, then Q_F is constant on $(F(x-), F(x))$; different treatments of quantiles within this interval are simply private randomization conditional on the state x.

Let $G \in \mathrm{MPC}(F)$ be bi-pooling. For a pooling interval $I = (a,b)$, let

$$B_I := (G(a), G(b-))$$

be its quantile block. The canonical decomposition implies that this block has length m_{F_I} and

$$Q_F(G(a)+r) = Q_{F_I}(r), \qquad 0 < r < m_{F_I}. \tag{9}$$

The blocks are disjoint, and quantiles outside them generate the residual law $G|_{K_G}$.

Proposition 5 (Implementing bi-pooling)**.** *Every bi-pooling $G \in \mathrm{MPC}(F)$ admits both:*

(i) a monotone quantile implementation; and

(ii) a pure quantile implementation in which, on every binary block, one interval is sent to the lower posterior mean and the two remaining tails are sent to the upper posterior mean.

In both constructions, residual quantiles are revealed and singleton blocks are pooled to their barycenters.

Taking our results together, for every upper semicontinuous objective u, there is an optimal distribution in (P) that admits both a monotone and a pure quantile implementation. We illustrate these two implementations in Figure 4.

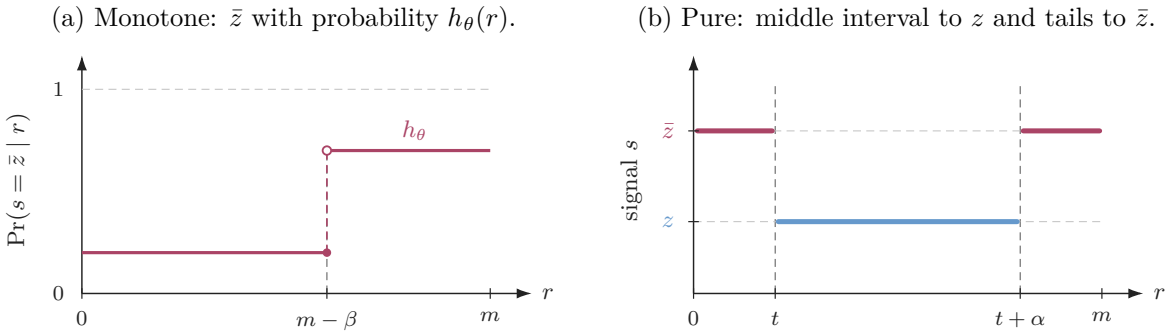

Figure 4: Two quantile implementations of the bi-pooling interval $G_I = \alpha\delta_z + \beta\delta_{\bar{z}}$.

Proof. Reveal $Q_F(q)$ outside the blocks. Fix a pooling interval I and write $m = m_{F_I}$, $Q = Q_{F_I}$, and $e = e_{F_I}$. If $G_I = m\delta_e$, send the whole block to e. It remains to consider

$$G_I = \alpha\delta_z + \beta\delta_{\bar{z}}, \qquad z < e < \bar{z}, \qquad \alpha + \beta = m.$$

For the *monotone implementation*, (8) gives $\int_{m-\beta}^{m} Q(r)\,\mathrm{d}r \geq \beta\bar{z}$. For $\theta \in [0,1]$, let

$$h_\theta(r) = (1-\theta)\frac{\beta}{m} + \theta\mathbf{1}\{r > m-\beta\}.$$

This function is nondecreasing and integrates to β. Moreover, $\int_0^m Q(r)h_\theta(r)\,\mathrm{d}r$ varies continuously in θ from βe to $\int_{m-\beta}^{m} Q(r)\,\mathrm{d}r$, so some θ satisfies $\int_0^m Q(r)h_\theta(r)\,\mathrm{d}r = \beta\bar{z}$. Then,

we construct the signal s at quantile $G(a) + r$ to send $\bar{z}$ with probability $h_\theta(r)$ and z otherwise. The two signals then have masses β and α and state moments $\beta\bar{z}$ and αz, respectively. Hence the block implements G_I. Since h_θ is nondecreasing and the pooling intervals are ordered, the resulting signal is monotone.

For the *pure implementation*, (8) gives

$$\int_0^\alpha Q(r)\,\mathrm{d}r \le \alpha z, \qquad \int_{m-\alpha}^m Q(r)\,\mathrm{d}r \ge \alpha z.$$

The map $t \mapsto \int_t^{t+\alpha} Q(r)\,\mathrm{d}r$ is continuous, so for some $t \in [0, m-\alpha]$ its value is αz. Send the interval $(t, t+\alpha)$ to z and the two remaining tails to $\bar{z}$. The first cell has mass α and state moment αz; its complement therefore has mass β and state moment $\beta\bar{z}$. This gives a pure implementation of G_I.

Applying either construction on every block and revealing all residual quantiles implements G. □

The two constructions also make explicit a familiar trade-off between monotonicity and determinism: the monotone implementation generally randomizes within a block, while the pure implementation need not be monotone.

4.4 Structured Objectives

The bi-pooling distributions feature a series of intervals in which the state is revealed and (bi-)pooled. We can derive further structure on this sequence based on the structure of the objective.

Proposition 6 (Piecewise convex and concave objectives)**.** *Let $u : [0,1] \to \mathbb{R}$ be upper semicontinuous, and let $0 = x_0 < x_1 < \cdots < x_k = 1$ be such that u is (weakly) convex or (weakly) concave on each piece $P_j = [x_{j-1}, x_j]$. A piece on which u is affine counts as convex. There is an optimal bi-pooling distribution G^* with the following properties:*

(i) no pooling interval is contained in a convex piece;

(ii) each concave piece contains at most one pooling interval, and that interval is pooled to the barycenter of its prior component;

(iii) every other pooling interval contains an interior switchpoint x_j.

If c of the pieces are concave, G^* has at most $c + k - 1$ pooling intervals, and hence at most $2k - 1$.

An analogous result formulated in terms of the dual price is given in Arieli et al. (2023, Proposition 3) for a nonatomic prior with full support. We directly bound the behavior of the optimal pooling intervals through a chain of modifications of a candidate distribution in the proof below.

Proof. Start from an optimal bi-pooling law G. If a pooling interval I is contained in a convex piece, replace G_I by F_I. This is feasible by Theorem 1, while convexity of u weakly raises the objective. Doing this for every such interval gives another optimal bi-pooling law, which we denote by $G^{(1)}$, and establishes part (i).

Now fix a concave piece P and consider the pooling intervals of $G^{(1)}$ contained in P. If there are none, do nothing. Otherwise, let a be the infimum of their left endpoints and b the supremum of their right endpoints. Let m denote the mass and e the barycenter of $G^{(1)}|_{(a,b)}$, and set

$$G^{(2)} = G^{(1)} - G^{(1)}|_{(a,b)} + m\delta_e.$$

Replacing $G^{(1)}|_{(a,b)}$ by its barycenter is a mean-preserving contraction, so $G^{(2)} \in \mathrm{MPC}(F)$. Concavity of u on P shows that the objective weakly increases.

It remains to identify the new pooling interval. The potential gap increases by $R_{G^{(1)}|_{(a,b)}} - m(\cdot - e)_+$. Let $\alpha = \inf \operatorname{supp}(G^{(1)}|_{(a,b)})$ and $\beta = \sup \operatorname{supp}(G^{(1)}|_{(a,b)})$. The correction vanishes outside $[a, b]$ and is positive on (α, β). If $x \in (a, \alpha]$, some original interval starts before x; it must contain x, since otherwise its posterior component would put $G^{(1)}|_{(a,b)}$-mass below α. Hence the old gap is positive at x. The same argument gives positivity on $[\beta, b)$. Since a and b are limits of binding endpoints, the new gap is zero there and positive throughout (a, b). Thus the contained intervals are replaced by the single pooling interval (a, b), whose posterior component is $m\delta_e$.

Repeat this operation on every concave piece. There is then at most one contained pooling interval per concave piece pooled to its barycenter, establishing (ii). Every remaining pooling interval contains a switchpoint (establishing (iii)), and disjoint pooling intervals contain distinct switchpoints. The number of pooling intervals is therefore at most $c + (k - 1)$. □

Proposition 6 reduces the optimization problem to a finite dimensional optimization over pooling intervals. Once an interval and its endpoint atom shares are chosen, its prior component F_I is determined. A contained concave interval is pooled to e_{F_I}; any other component has one or two support points, with the weights and feasibility of a proposed pair given by the requirement that $G_I \in \mathrm{MPC}(F_I)$. Thus a candidate is described by finitely many interval endpoints, atom shares, and posterior means.

S-shaped Objectives and Generalized Censorship An important class of objective functions are those with just one switching point, the (inverse) S-shaped objectives studied in Kolotilin et al. (2022). They show that if the prior admits a density, there exists an optimal policy that is upper/lower censorship: it reveals the state below/above a cutoff q and pools all states above/below the cutoff. For a discrete prior, Alonso and Câmara (2016); Kolotilin et al. (2024) establish an analogous stochastic censorship result.

The censorship result generalizes to general priors, suitably adjusting the definition of a censorship policy to allow for the splitting of atoms, as follows.

Definition 4 (Generalized censorship)**.** A distribution $G \in \mathrm{MPC}(F)$ is a *generalized lower-censorship distribution* if either $G = F$, or G has a unique pooling interval $I = (\min \operatorname{supp}(F), a)$ for some $a \leq \max \operatorname{supp}(F)$, and

$$G_I = m_{F_I} \delta_{e_{F_I}}, \qquad \Delta G(\min \operatorname{supp}(F)) = 0.$$

A distribution $G \in \mathrm{MPC}(F)$ is a *generalized upper-censorship distribution* if either $G = F$, or G has a unique pooling interval $I = (b, \max \operatorname{supp}(F))$ for some $b \geq \min \operatorname{supp}(F)$, and

$$G_I = m_{F_I} \delta_{e_{F_I}}, \qquad \Delta G(\max \operatorname{supp}(F)) = 0.$$

To see that the definition directly extends the censorship notion, let us parse it, focusing on lower censorship for concreteness: All states below the cutoff a are pooled. This holds for the pooling interval $I = (\min \operatorname{supp}(F), a)$ as G_I is a point mass. If F has a mass point at the minimum of its support, all of this mass is also pooled into the interval, as ensured by the condition that $\Delta G(\min \operatorname{supp}(F)) = 0$. As I is the only pooling interval, all states above a are revealed. If F has a point mass at a, the definition does not restrict how this mass is split between the pooling interval and being revealed.

Proposition 7 (Optimal censorship)**.** *Let $u \in C^1[0,1]$ be concave on $[0,c]$ and convex on $[c,1]$ for some $c \in [0,1]$. For every prior F, a generalized lower-censorship distribution maximizes $\int u \, \mathrm{d}G$ over $G \in \mathrm{MPC}(F)$.*
If the order of convexity and concavity is reversed, a generalized upper-censorship distribution maximizes. If the concavity and convexity assumptions are strict, the maximizing distribution is unique.

Our proof follows the template of Kolotilin et al. (2022), generalized to quantile space to allow for atoms and for an objective u that is merely once continuously differentiable, and is relegated to the appendix.

5 Conclusion

This paper shows that the tools and insights in the one-dimensional mean-dependent persuasion problem generalize to the case of a general distribution F in a useful yet rich way.

An analogous generalization in other domains, such as the higher-dimensional case (Kleiner et al., forthcoming) or convex function intervals (Augias and Uhe, 2025) is left for future research.

References

Akbarpour, Mohammad, Eric Budish, Piotr Dworczak, and Scott Duke Kominers, "An Economic Framework for Vaccine Prioritization," *The Quarterly Journal of Economics*, February 2024, *139* (1), 359–417.

_ , Piotr Dworczak, and Scott Duke Kominers, "Redistributive Allocation Mechanisms," *Journal of Political Economy*, June 2024, *132* (6), 1831–1875.

Alonso, Ricardo and Odilon Câmara, "Political Disagreement and Information in Elections," *Games and Economic Behavior*, November 2016, *100*, 390–412.

Arieli, Itai, Yakov Babichenko, Rann Smorodinsky, and Takuro Yamashita, "Optimal Persuasion via Bi-Pooling," *Theoretical Economics*, 2023, *18* (1), 15–36.

Augias, Victor and Lina Uhe, "The Economics of Convex Function Intervals," Working Paper, arXiv:2510.20907, October 2025.

Bergemann, Dirk, Tibor Heumann, and Stephen Morris, "Information Design and Mechanism Design: An Integrated Framework," Cowles Foundation Discussion Paper 2494, arXiv:2601.17267, January 2026.

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

_ and Giorgio Martini, "The Simple Economics of Optimal Persuasion," *Journal of Political Economy*, October 2019, *127* (5), 1993–2048.

Gentzkow, Matthew and Emir Kamenica, "A Rothschild-Stiglitz Approach to Bayesian Persuasion," *American Economic Review*, May 2016, *106* (5), 597–601.

Goldstein, Itay and Yaron Leitner, "Stress Tests and Information Disclosure," *Journal of Economic Theory*, September 2018, *177*, 34–69.

Kamenica, Emir and Matthew Gentzkow, "Bayesian Persuasion," *American Economic Review*, October 2011, *101* (6), 2590–2615.

Kleiner, Andreas, Benny Moldovanu, and Philipp Strack, "Extreme Points and Majorization: Economic Applications," *Econometrica*, 2021, *89* (4), 1557–1593.

_ , _ , _ , and Mark Whitmeyer, "The Extreme Points of Mean-Preserving Contractions," *Journal of Economic Theory*, forthcoming.

Kolotilin, Anton and Andriy Zapechelnyuk, "Persuasion Meets Delegation," *Econometrica*, 2025, *93* (1), 195–228.

_ , **Hongyi Li, and Andy Zapechelnyuk**, "On Monotone Persuasion," Working Paper, arXiv:2412.14400, December 2024.

_ , **Timofiy Mylovanov, and Andriy Zapechelnyuk**, "Censorship as Optimal Persuasion," *Theoretical Economics*, 2022, *17* (2), 561–585.

Ostrizek, Franz, "Vague by Design: Performance Evaluation and Learning From Wages," *mimeo*, 2024.

Ostrovsky, Michael and Michael Schwarz, "Information Disclosure and Unraveling in Matching Markets," *American Economic Journal: Microeconomics*, May 2010, *2* (2), 34–63.

Strassen, V., "The Existence of Probability Measures with Given Marginals," *The Annals of Mathematical Statistics*, 1965, *36* (2), 423–439.

Xiao, Peiran, "Allocating Positional Goods: A Mechanism Design Approach," Working Paper, arXiv:2411.06285, July 2026.

A Proofs

Proof of Lemma 1: (i) For $y \le x$ we have $(x-y)_+ = \int_0^x \mathbf{1}\{t \ge y\}\,\mathrm{d}t$, and for $y > x$ both sides vanish, so by Tonelli

$$R_H(x) = \int_{[0,1]} \int_0^x \mathbf{1}\{t \ge y\}\,\mathrm{d}t\,\mathrm{d}H(y) = \int_0^x H([0,t])\,\mathrm{d}t.$$

The integrand $H(t)$ is nondecreasing and bounded by m_H, giving convexity, monotonicity, and the Lipschitz bound; $R_H(0) = 0$ is immediate. (ii) $R_H(1) = \int (1-y)\,\mathrm{d}H(y) = m_H - \int y\,\mathrm{d}H(y)$. (iii) By (i), $H(t)$ is, for almost every t, the derivative of R_H at t; since $t \mapsto H(t)$ is right-continuous on $[0,1)$, it is determined at every $t \in [0,1)$ by its values on a co-null set. This pins down H on all sets $[0,t]$, $t < 1$, hence on the Borel σ-algebra of $[0,1)$; the atom at 1 is then $m_H - H(1-)$. For a signed measure, apply the argument to the Jordan decomposition of the cdf: the cdf is still right-continuous of bounded variation and is a.e. the derivative of the potential. □

Proof of Lemma 2: If the integral inequality holds, applying it to the constant function, the identity function, and their negatives implies equal mass and first moment. Applying it to $\varphi_x(y) = (x-y)_+$ for each x gives $R_\nu \le R_\mu$; equal mass and first moment give $R_\nu(1) = R_\mu(1)$.

Conversely, suppose $\nu \in \mathrm{MPC}(\mu)$. Equal mass and first moment imply, for every $t \in [0,1]$,

$$\int (y-t)_+\,\mathrm{d}\mu(y) - \int (y-t)_+\,\mathrm{d}\nu(y) = R_\mu(t) - R_\nu(t) \ge 0.$$

Every piecewise-affine convex function has the form

$$a + by + \sum_{j=1}^{k} c_j (y - t_j)_+, \qquad c_j \ge 0,$$

so the desired inequality holds for all such functions. Finally, the piecewise-affine interpolants of any continuous convex φ on successively finer grids are convex and converge uniformly to φ. Passing to the limit establishes the result. □

Proof of Lemma 3: Nonemptiness: $F \in \mathrm{MPC}(F)$. Convexity: $G \mapsto R_G(x)$ is affine in G for each fixed x, and the defining conditions are an inequality and an equality in these affine functionals. Compactness: The set of measures with unit mass in $\Delta([0,1])$ is weak* compact by Prokhorov's theorem, so it suffices to show that $\mathrm{MPC}(F)$ is weak* closed. To see this, note that for each fixed x the function $y \mapsto (x-y)_+$ is continuous and bounded. Hence, if $G_n \to G$, then $R_{G_n}(x) \to R_G(x)$. Pointwise limits preserve the inequality $R_{G_n}(x) \le R_F(x)$ and the equality at $x = 1$. □

Before proceeding, we state the following lemma that formalizes intermediate steps in the construction towards Theorem 1.

Lemma 5 (Endpoint slopes)**.** *Let $x \in K_G$.*

(i) If $x < 1$, then $F(x) \geq G(x)$; if $x > 0$, then $G(x-) \geq F(x-)$.

(ii) Consequently, for every $x \in K_G \cap (0,1)$, the two nonnegative numbers $D'_G(x+) = F(x) - G(x)$ and $-D'_G(x-) = G(x-) - F(x-)$ satisfy

$$D'_G(x+) - D'_G(x-) = \Big(F(x) - G(x)\Big) + \Big(G(x-) - F(x-)\Big) = \Delta F(x) - \Delta G(x),$$

and each lies in $[0, \Delta F(x)]$. At $x = 0$, $D'_G(0+) \in [0, \Delta F(0)]$. At $x = 1$, $-D'_G(1-) \in [0, \Delta F(1)]$.

Note that this lemma also establishes the statement about the λ_I^a, λ_I^b in the main text.

Proof of Lemma 5: (i) At a binding point $x < 1$ we have $D_G(x) = 0$ and $D_G \geq 0$ on $[x, 1]$, so the right derivative in (1) is nonnegative: $F(x) \geq G(x)$. Symmetrically, for $x > 0$, $D_G \geq 0$ on $[0, x]$ with $D_G(x) = 0$ forces the left derivative to be nonpositive: $G(x-) \geq F(x-)$.

(ii) The identity is algebra:

$$\Big(F(x)-G(x)\Big)+\Big(G(x-)-F(x-)\Big) = \Big(F(x)-F(x-)\Big)-\Big(G(x)-G(x-)\Big) = \Delta F(x)-\Delta G(x).$$

Both summands are nonnegative by (i), so each is at most $\Delta F(x) - \Delta G(x) \leq \Delta F(x)$. At $x = 0$, $F(0-) = G(0-) = 0$ gives $F(0) - G(0) \leq F(0) = \Delta F(0)$. At $x = 1$, $G(1-) - F(1-) \leq F(1) - F(1-) = \Delta F(1)$ since $G(1-) \leq 1 = F(1)$. □

Proof of Theorem 1: Fix a pooling interval $I = (a, b)$.

Step 1: cdf identity. For $t \in [a, b)$,

$$F_I(t) = \lambda_I^a + F((a,t]) = \Big(F(a)-G(a)\Big)+\Big(F(t)-F(a)\Big) = F(t)-G(a), \qquad G_I(t) = G(t)-G(a),$$

so

$$F_I(t) - G_I(t) = F(t) - G(t), \qquad t \in [a, b). \tag{10}$$

For $t < a$ both cdfs vanish, and for $t \geq b$ both are constant.

Step 2: equal mass and moment. Using (10) and integrating,

$$R_{F_I}(x) - R_{G_I}(x) = \int_a^x \Big(F(t) - G(t)\Big)\,\mathrm{d}t = D_G(x) - D_G(a) = D_G(x) \qquad \text{for } x \in [a, b].$$

In particular $R_{F_I} - R_{G_I} \geq 0$ on $[a, b]$, with value 0 at both a and b. The masses agree:

$$m_{F_I} = \lambda_I^a + F((a,b)) + \lambda_I^b = \Big(F(a)-G(a)\Big)+\Big(F(b-)-F(a)\Big)+\Big(G(b-)-F(b-)\Big) = G(b-)-G(a) = m_{G_I}.$$

Since both components are supported in $[a, b]$ and have equal mass, their potentials on $[b, 1]$ differ by the difference of first moments times 1; vanishing of the difference at $x = b$, where $R_H(b) = m_H b - \int y \, \mathrm{d}H(y)$ for any H supported in $[a, b]$, forces the first moments to agree, and then the potential difference vanishes identically on $[b, 1]$. It also vanishes on $[0, a]$. Hence

$$R_{F_I}(x) - R_{G_I}(x) = \begin{cases} D_G(x), & x \in [a, b], \\ 0, & \text{otherwise}, \end{cases} \tag{11}$$

which proves $G_I \in \mathrm{MPC}(F_I)$. The common mass is positive: if $m_{G_I} = 0$ then (11) would force $D_G \equiv 0$ on I, contradicting $D_G > 0$ there.

Step 3: the residual identity. Set $\sigma_I = F_I - G_I$, a signed measure of total mass zero, and

$$\rho = (F - G) - \sum_I \sigma_I.$$

The sum converges in total variation because $\sum_I |\sigma_I|([0, 1]) \le \sum_I (m_{F_I} + m_{G_I}) \le 2$. By (11), for each x at most one term of $\sum_I R_{\sigma_I}(x)$ is nonzero, and the sum equals $D_G(x)$ when $x \in \overline{I}$ for some pooling I and 0 when $x \in K_G$; in both cases $\sum_I R_{\sigma_I}(x) = D_G(x)$. Hence $R_\rho \equiv 0$ on $[0, 1]$, and ρ has total mass zero because $F - G$ and every σ_I do. Lemma 1(iii) for signed measures gives $\rho = 0$, i.e.

$$F - \sum_I F_I = G - \sum_I G_I.$$

Finally, the measures $G_I = G|_I$ are restrictions of G to the pairwise disjoint open sets I, whose union is U_G; hence $G - \sum_I G_I = G|_{K_G} \ge 0$. □

Proof of Corollary 1. Nonnegativity is Lemma 5 part (i). To see that $\lambda^x_{I_{x-}}$ can be taken zero if no pooling interval touches x from the left, note that then there are binding points $x_n \nearrow x$. Hence

$$D_G(x_n) = D_G(x) = 0 \quad \Longrightarrow \quad D'_G(x-) = 0,$$

and therefore $G(x-) - F(x-) = -D'_G(x-) = 0$. The right side is analogous: $F(x) - G(x) = D'_G(x+) = 0$ whenever no pooling interval touches x from the right.

The identity then follows since

$$\lambda^x_{I_{x-}} + \lambda^x_{I_{x+}} = \big(G(x-) - F(x-)\big) + \big(F(x) - G(x)\big) = \Delta F(x) - \Delta G(x). \qquad \square$$

Proof of Lemma 4: Let $\ell = \min \mathrm{supp}(F)$ and $r = \max \mathrm{supp}(F)$ be the endpoints of the prior's support. Every mean-preserving contraction of F is supported on $[\ell, r]$. If $\ell = r$, the only feasible distribution is δ_ℓ, so the claim is immediate. Henceforth we assume $\ell < r$.

Adding a constant to u does not change its maximizers, so we may assume $u \geq 0$. For $v \in C[0,1]$, write

$$V(v) = \max_{J \in \mathrm{MPC}(F)} \int v \, \mathrm{d}J, \qquad V = V(u).$$

These maxima exist by Lemma 3. We will construct a convex function $p \geq u$ with $\int p \, \mathrm{d}F = V$.

Choose Lipschitz functions u_n such that $u \leq u_n \leq u + \frac{1}{n}$. Applied on $[\ell, r]$, after an affine rescaling, Theorem 2 in Dworczak and Martini (2019) provides continuous convex functions $p_n : [\ell, r] \to \mathbb{R}$ satisfying $p_n \geq u_n$ and

$$\int p_n \, \mathrm{d}F = V(u_n).$$

The approximation changes every feasible payoff by at most $1/n$, so

$$p_n \geq u \geq 0, \qquad V \leq \int p_n \, \mathrm{d}F \leq V + \frac{1}{n}. \tag{12}$$

We next pass to a limit of these prices. First, fix $0 < \delta < (r-\ell)/4$. The neighborhoods $[\ell, \ell+\delta]$ and $[r-\delta, r]$ both have positive F-mass. Since $p_n \geq 0$ and $\int p_n \, \mathrm{d}F \leq V + 1$, each neighborhood contains a point at which p_n is bounded by a constant independent of n. The chord joining these two points bounds p_n from above on $[\ell+\delta, r-\delta]$. Together with $p_n \geq 0$, convexity then bounds the slopes on $[\ell + 2\delta, r - 2\delta]$. The prices are therefore uniformly bounded and equicontinuous on every closed subinterval of (ℓ, r). By Arzelà–Ascoli and a diagonal argument, we may pass to a subsequence along which p_n converges locally uniformly to a finite convex function p on (ℓ, r).

Define the endpoint values by $p(\ell) = \lim_{x \searrow \ell} p(x)$ and $p(r) = \lim_{x \nearrow r} p(x)$. These limits exist in $[0, +\infty]$ by the convexity and nonnegativity of p_n. The resulting function on $[\ell, r]$ is lower semicontinuous and convex, and continuity of u ensures $p \geq u$ also at the endpoints. Moreover, $p \leq \liminf_n p_n$ on the whole of $[\ell, r]$. In the interior this follows from convergence. At the left endpoint, fix $\ell < x < z < r$. The convexity inequality for p_n gives

$$\liminf_n p_n(\ell) \geq \frac{z-\ell}{z-x} p(x) - \frac{x-\ell}{z-x} p(z).$$

Letting $x \searrow \ell$ proves the claim at ℓ; the argument at r is symmetric. Fatou's lemma and (12) now give

$$\int p \, \mathrm{d}F \leq \liminf_n \int p_n \, \mathrm{d}F = V.$$

To obtain the reverse inequality, note that using $p \geq u$ and the convex order inequality, which extends to the lower semicontinuous p by monotone convergence, yields

$$V = \int u \, \mathrm{d}G \leq \int p \, \mathrm{d}G \leq \int p \, \mathrm{d}F \leq V.$$

Thus $\int p\,\mathrm{d}F = \int p\,\mathrm{d}G = V$, and p is integrable under every feasible distribution. Furthermore, $\int (p-u)\,\mathrm{d}G = 0$. The difference $p-u$ is nonnegative and lower semicontinuous, so its zero set is closed. It follows that

$$\operatorname{supp}(G) \subseteq \{x \in [\ell, r] : p(x) = u(x)\}. \tag{13}$$

Now let H satisfy the two conditions in the lemma. By the support condition and (13), we already have $\int u\,\mathrm{d}H = \int p\,\mathrm{d}H$. To use the binding condition, we express differences in expected prices in terms of potential differences.

Because p is convex, its right derivative $p'(x+)$ is nondecreasing in x. Its increases define a positive Radon measure $\mu = D^2 p$ on (ℓ, r), given by $\mu((s,t]) = p'(t+) - p'(s+)$ for $\ell < s < t < r$. Fix $x_0 \in (\ell, r)$. Integrating the derivative gives

$$\begin{aligned} p(x) =& p(x_0) + p'(x_0+)(x - x_0) \\ &+ \int_{(\ell, x_0]} (t - x)_+ \,\mathrm{d}\mu(t) + \int_{(x_0, r)} (x - t)_+ \,\mathrm{d}\mu(t), \qquad x \in [\ell, r]. \end{aligned}$$

The formula extends to the endpoints by monotone convergence. The hinge terms are nonnegative, so Tonelli allows us to integrate this representation against F and any feasible J. Each resulting integral is finite because $\int p\,\mathrm{d}F$ and $\int p\,\mathrm{d}J$ are finite. Subtracting the two formulas cancels the affine terms, since F and J have the same mass and mean. Using $(x-t)_+ = (x-t) + (t-x)_+$ for the remaining terms gives[4]

$$\int p\,\mathrm{d}F - \int p\,\mathrm{d}J = \int_{(\ell, r)} \big(R_F(t) - R_J(t)\big)\,\mathrm{d}\mu(t). \tag{14}$$

Taking $J = G$ in (14) gives $\int D_G\,\mathrm{d}\mu = 0$. Since $D_G = R_F - R_G$ is nonnegative and continuous, μ assigns no mass to $U_G = \{D_G > 0\}$. Thus μ is carried by $K_G \cap (\ell, r)$.

Then, since $\operatorname{supp}(H) \subseteq \operatorname{supp}(G)$, (13) implies that $\int u\,\mathrm{d}H = \int p\,\mathrm{d}H$. Since $R_H = R_F$ on K_G the right side of (14) vanishes also for $J = H$. Therefore,

$$\int u\,\mathrm{d}H = \int p\,\mathrm{d}H = \int p\,\mathrm{d}F = V,$$

which proves the optimality of H. □

Proof of Proposition 7. We prove upper censorship; the lower-censorship result is exactly analogous.

If $c \geq \max \operatorname{supp}(F)$, full disclosure is optimal by convexity. If $c \leq \min \operatorname{supp}(F)$, no disclosure is optimal by concavity. Both are generalized censorship policies. Hence suppose

[4]Note that this does not require $\mu((\ell, r)) < \infty$. See also Appendix A.4, Lemma 2 in Dworczak and Martini (2019).

$c \in (\min \operatorname{supp}(F), \max \operatorname{supp}(F))$. We can hence normalize to let $\min \operatorname{supp}(F) = 0$ and $\max \operatorname{supp}(F) = 1$ for the remainder of this proof, to lighten notation.

For $q \in [0, 1)$, let

$$m(q) = \frac{1}{1-q} \int_q^1 Q_F(s) \, \mathrm{d}s$$

be the mean of the upper $1 - q$ quantiles. The policy that reveals $Q_F(s)$ for $s < q$ and pools at $m(q)$ for $s > q$ has payoff

$$W(q) = \int_0^q u(Q_F(s)) \, \mathrm{d}s + (1-q)u(m(q)).$$

This function extends continuously to $q = 1$, where the policy is full disclosure, and is absolutely continuous on $[0, 1]$. Both $Q_F(q)$ and $m(q)$ are nondecreasing, and $Q_F(q) \leq m(q)$.

For $y \in [0, 1]$, let

$$L_y(x) = u(y) + u'(y)(x - y)$$

be the tangent line to u at y. Direct differentiation gives, for almost every $q < 1$,

$$W'(q) = u(Q_F(q)) - L_{m(q)}(Q_F(q)).$$

Thus W increases when u lies above the tangent at the cutoff state, and falls when it lies below. If $m(q) \leq c$, then $Q_F(q) \leq m(q) \leq c$, so convexity gives $W'(q) \geq 0$.

We next show that, on the region where $m(q) \geq c$, once W' becomes nonpositive, it remains nonpositive. Let $q_1 < q_2$ be differentiability points of W with $m(q_1) \geq c$, and suppose $W'(q_1) \leq 0$. Write $x_i = Q_F(q_i)$ and $y_i = m(q_i)$. Then $x_1 \leq x_2$ and $y_1 \leq y_2$. Concavity on $[c, 1]$ implies

$$L_{y_2}(x) \geq L_{y_1}(x) \qquad \text{for every } x \leq y_1.$$

Hence $u(x_1) \leq L_{y_2}(x_1)$. If $x_2 \geq c$, concavity gives $u(x_2) \leq L_{y_2}(x_2)$. If $x_2 < c$, the function $u - L_{y_2}$ is convex on $[0, c]$ and nonpositive at both x_1 and c, so it is nonpositive at x_2 as well. Thus $W'(q_2) \leq 0$.

It follows that W is nondecreasing and then nonincreasing. It is nondecreasing while $m(q) < c$, and nonincreasing once $q > F(c)$. At $q = F(c)$, all $s > q$ satisfy $Q_F(s) \geq c$, so $m(F(c)) \geq c$. Therefore some maximizer q satisfies

$$m(q) \geq c, \qquad q \leq F(c).$$

In particular, $q < 1$.

Write $m = m(q)$ and $L = L_m$. Suppose first that $q > 0$, and let $x_- = Q_F(q-)$ and $x_+ = Q_F(q+)$. At a maximum, the left derivative of W is nonnegative and the right

derivative is nonpositive. The one-sided version of the derivative formula therefore gives

$$u(x_-) \geq L(x_-), \qquad u(x_+) \leq L(x_+).$$

Also, $q \leq F(c)$ implies $x_- \leq c$, while $m \geq c$ and concavity imply $u(c) \leq L(c)$.

Let $h = u - L$, and let a be the first point in $[x_-, c]$ at which $h \leq 0$. Then $h(a) = 0$ and $x_- \leq a \leq x_+$. Convexity of h on $[0, c]$ gives $h \leq 0$ on $[a, c]$, and concavity of u gives $h \leq 0$ on $[c, 1]$. Since h is nonpositive immediately to the right of a, we have $u'(a) \leq u'(m)$. Therefore

$$p(x) = \begin{cases} u(x), & x \leq a, \\ L(x), & x \geq a \end{cases}$$

is convex and satisfies $p \geq u$.

If $q = 0$, the right derivative of W gives $u(0) \leq L(0)$. We also have $u(c) \leq L(c)$. Since $u - L$ is convex on $[0, c]$ and nonpositive at both endpoints, and $L \geq u$ on $[c, 1]$, we may take $a = 0$ and $p = L$.

For almost every $s < q$, $Q_F(s) \leq a$, so $p(Q_F(s)) = u(Q_F(s))$. For almost every $s > q$, $Q_F(s) \geq a$, so $p(Q_F(s)) = L(Q_F(s))$. Since the upper tail has mean m and L is affine,

$$\int p \, \mathrm{d}F = \int_0^q u(Q_F(s)) \, \mathrm{d}s + (1 - q)L(m) = W(q).$$

Let G be the distribution induced by this policy. Then $\int u \, \mathrm{d}G = W(q)$. For every $H \in \mathrm{MPC}(F)$,

$$\int u \, \mathrm{d}H \leq \int p \, \mathrm{d}H \leq \int p \, \mathrm{d}F = \int u \, \mathrm{d}G, \tag{15}$$

where the second inequality follows from convex order. Thus G is optimal.

If Q_F is constant on $(q, 1)$, pooling changes nothing (as we are pooling quantiles within an atom) and $G = F$. Otherwise, let $b = Q_F(q+)$. Then, $(b, 1)$ is the unique pooling interval, and its posterior component is concentrated at m. All prior mass at 1 belongs to the pooled component; since it is nondegenerate, $m < 1$ and $\Delta G(1) = 0$. Therefore G is a generalized upper-censorship distribution.

Suppose now that u is strictly convex on $[0, c]$ and strictly concave on $[c, 1]$, and let $H \in \mathrm{MPC}(F)$ be optimal. If $c \geq \max \operatorname{supp}(F)$ or $c \leq \min \operatorname{supp}(F)$, strict Jensen's inequality gives $H = F$ or $H = \delta_{e_F}$, respectively. In the remaining case, both inequalities in (15) must be equalities. Strictness implies

$$\{x : p(x) = u(x)\} \subseteq [0, a] \cup \{m\},$$

since $h = u - L$ is strictly convex on $[0, c]$ with $h(a) \leq 0$ and $h(c) \leq 0$, so $h < 0$ on (a, c), and strictly concave on $[c, 1]$ with $h(m) = h'(m) = 0$ and $m \geq c$, so $h < 0$ on $[c, 1] \backslash \{m\}$. Hence H is supported on this set.

Moreover, strict convexity of $p = u$ on $[0, a]$ means that its curvature measure charges every open subinterval of $(0, a)$. The curvature identity (14), which holds for any convex p with finite integrals against F and H, therefore implies $R_H = R_F$ on $(0, a)$, since their difference is continuous and nonnegative. Consequently, $H = F = G$ on $[0, a)$. The remaining masses of H and G are supported on $\{a, m\}$ and are uniquely determined by total mass and first moment. Thus $H = G$. □

B Calculations for Examples and Additional Material

B.1 The split-atom example

This section provides detailed calculations for Example 1, showing that not every bi-pooling distribution is extreme. It then uses this example to show that extreme points of mean-preserving contractions are not preserved by weak limits.

Recall that

$$F = \tfrac{1}{3}\delta_0 + \tfrac{1}{3}\delta_{1/2} + \tfrac{1}{3}\delta_1, \qquad G_t = \frac{4-t}{9}\,\delta_{1/10} + \frac{4t-1}{9}\,\delta_{2/5} + \frac{8-9t}{15}\,\delta_{7/10} + \frac{2+4t}{15}\,\delta_{19/20},$$

with $t \in [1/4, 8/9]$; for $t \in (1/4, 8/9)$ all four weights are positive. The parameter t is the share of the atom at $1/2$ assigned to the left. Write $I_1 = (0, 1/2)$ and $I_2 = (1/2, 1)$. With $\lambda_{I_1}^{1/2} = t/3$ and $\lambda_{I_2}^{1/2} = (1-t)/3$, the components of (2) are

$$F_{I_1} = \tfrac{1}{3}\delta_0 + \tfrac{t}{3}\delta_{1/2}, \qquad F_{I_2} = \tfrac{1-t}{3}\delta_{1/2} + \tfrac{1}{3}\delta_1,$$

and G_{I_1}, G_{I_2} are the restrictions of G_t to $\{1/10, 2/5\}$ and $\{7/10, 19/20\}$. Direct calculation gives

	mass	first moment
F_{I_1}, G_{I_1}	$(1+t)/3$	$t/6$
F_{I_2}, G_{I_2}	$(2-t)/3$	$(3-t)/6$

so each posterior component matches its prior component in mass and mean, as in Theorem 1(i).

On $[0, 1/2]$, $D_{G_t} = R_{F_{I_1}} - R_{G_{I_1}}$. It is affine between $0, 1/10, 2/5, 1/2$, with respective values

$$0, \quad \frac{1}{30}, \quad \frac{t}{30}, \quad 0.$$

On $[1/2, 1]$, $D_{G_t} = R_{F_{I_2}} - R_{G_{I_2}}$. It is affine between $1/2, 7/10, 19/20, 1$, with respective values

$$0, \quad \frac{1-t}{15}, \quad \frac{1}{60}, \quad 0.$$

For $t \in (1/4, 8/9)$ all posterior weights are positive, so these values show that the pooling intervals are exactly I_1 and I_2, each containing two posterior atoms, and that the shares

above are the ones given by (1); they exhaust the atom at 1/2. Finally, every weight in G_t is affine in t. Hence, for sufficiently small $\varepsilon > 0$,

$$G_t = \tfrac{1}{2}G_{t-\varepsilon} + \tfrac{1}{2}G_{t+\varepsilon},$$

with two distinct feasible distributions on the right. Thus G_t is not extreme and therefore not exposed.

Example 3 (Extremality need not survive weak* convergence of the prior)**.** Recall that $\mathsf{U}[a, b]$ denotes the uniform distribution on $[a, b]$, and, for $0 < \epsilon < \frac{3}{25}$, set

$$\begin{aligned} F_\epsilon =& \frac{1}{3}\mathsf{U}[0,\epsilon] + \frac{1}{6}\mathsf{U}\left[\frac{1}{2} - \frac{3\epsilon}{2}, \frac{1}{2} - \frac{\epsilon}{2}\right] \\ &+ \frac{1}{6}\mathsf{U}\left[\frac{1}{2} + \frac{\epsilon}{2}, \frac{1}{2} + \frac{3\epsilon}{2}\right] + \frac{1}{3}\mathsf{U}[1-\epsilon, 1]. \end{aligned}$$

Then F_ϵ is nonatomic and

$$F_\epsilon \to F = \frac{1}{3}\delta_0 + \frac{1}{3}\delta_{1/2} + \frac{1}{3}\delta_1,$$

the prior of Example 1. Consider from the previous example

$$G_{1/2} = \frac{7}{18}\delta_{1/10} + \frac{1}{9}\delta_{2/5} + \frac{7}{30}\delta_{7/10} + \frac{4}{15}\delta_{19/20}.$$

A direct potential calculation gives that $\{R_{F_\epsilon} - R_{G_{1/2}} > 0\} = \left(0, \frac{1}{2} - \frac{\epsilon}{2}\right) \cup \left(\frac{1}{2} + \frac{\epsilon}{2}, 1\right)$, so $G_{1/2}$ is bi-pooling relative to F_ϵ and $G_{1/2} \in \mathrm{MPC}(F_\epsilon)$ for all $\epsilon < \frac{3}{25}$. Since F_ϵ is nonatomic, Corollary 2 implies

$$G_{1/2} \in \mathrm{ext}(\mathrm{MPC}(F_\epsilon)).$$

In the limit, however, we can split

$$G_{1/2} = \frac{1}{2}G_{1/2-\eta} + \frac{1}{2}G_{1/2+\eta}$$

for sufficiently small $\eta > 0$, with both distributions on the right distinct and a mean-preserving contraction of F. Thus

$$G_{1/2} \notin \mathrm{ext}(\mathrm{MPC}(F)).$$

Hence extremality need not survive weak* convergence of the prior. In our example, the support gap around 1/2 becomes an atom in the limit, and its split between the two adjacent pooling regions creates the new degree of freedom.

B.2 The one-way-gate example

This section provides the details and verification for Example 2, showing that not every extreme point is exposed.

Both F and G are probability measures with mean $1/2$. A direct calculation gives

$$D_G(x) = R_F(x) - R_G(x) = \begin{cases} \dfrac{x^2}{2}, & 0 \le x \le \dfrac{1}{6}, \\ \dfrac{1}{2}\left(x - \dfrac{1}{4}\right)^2 + \dfrac{1}{96}, & \dfrac{1}{6} \le x \le \dfrac{1}{3}, \\ \dfrac{1}{2}\left(x - \dfrac{1}{2}\right)^2, & \dfrac{1}{3} \le x \le \dfrac{1}{2}, \\ 0, & \dfrac{1}{2} \le x \le 1. \end{cases}$$

Hence $D_G \ge 0$, with

$$D_G > 0 \quad \text{on } (0, 1/2), \qquad D_G = 0 \quad \text{on } \{0\} \cup [1/2, 1].$$

Thus $G \in \mathrm{MPC}(F)$, $U_G = (0, 1/2)$, $K_G = \{0\} \cup [1/2, 1]$ and G is bi-pooling because its posterior component on the unique pooling interval is $G|_{(0,1/2)} = \frac{1}{4}\delta_{1/6} + \frac{1}{4}\delta_{1/3}$.

Now define the signed measure $\xi = -\delta_{1/6} + 2\delta_{1/3} - \delta_{1/2}$. It has zero total mass and zero first moment. Its potential is

$$R_\xi(x) = \begin{cases} 0, & 0 \le x \le \dfrac{1}{6}, \\ \dfrac{1}{6} - x, & \dfrac{1}{6} \le x \le \dfrac{1}{3}, \\ x - \dfrac{1}{2}, & \dfrac{1}{3} \le x \le \dfrac{1}{2}, \\ 0, & \dfrac{1}{2} \le x \le 1. \end{cases}$$

In particular, $R_\xi \le 0$ and $R_\xi = 0$ on K_G. We claim that

$$\{H \in \mathrm{MPC}(F) : \mathrm{supp}(H) \subseteq \mathrm{supp}(G),\ R_H = R_F \text{ on } K_G\} = \left\{G + t\xi : 0 \le t \le \frac{1}{4}\right\}. \quad (16)$$

First, let $0 \le t \le 1/4$. Then

$$G + t\xi = \left(\frac{1}{4} - t\right)\delta_{1/6} + \left(\frac{1}{4} + 2t\right)\delta_{1/3} + \left(\frac{1}{4} - t\right)\delta_{1/2} + \frac{1}{4}\delta_1$$

so $G + t\xi$ is a positive probability measure with the same mean as F. Moreover,

$$D_{G+t\xi} = R_F - R_{G+t\xi} = D_G - tR_\xi \ge 0.$$

Thus $G + t\xi \in \mathrm{MPC}(F)$. Since $R_\xi = 0$ on K_G, $R_{G+t\xi} = R_G = R_F$ on K_G and its support is contained in $\mathrm{supp}(G)$.

Conversely, let $H \in \mathrm{MPC}(F)$ satisfy

$$\mathrm{supp}(H) \subseteq \mathrm{supp}(G), \qquad R_H = R_F \quad \text{on } K_G.$$

Write

$$H - G = a\delta_{1/6} + b\delta_{1/3} + c\delta_{1/2} + d\delta_1.$$

Because $R_H = R_G$ on $[1/2, 1]$, the affine function

$$R_{H-G}(x) = a\left(x - \frac{1}{6}\right) + b\left(x - \frac{1}{3}\right) + c\left(x - \frac{1}{2}\right), \qquad \frac{1}{2} < x < 1,$$

vanishes identically. Hence

$$a + b + c = 0, \qquad \frac{a}{6} + \frac{b}{3} + \frac{c}{2} = 0.$$

Solving the two remaining equations gives $b = -2a$ and $c = a$. Since $H - G$ also has total mass zero, we have $d = 0$. Therefore $H = G - a\xi$. Setting $t = -a$, feasibility requires $t \geq 0$. Indeed, if $t < 0$, then for $0 < h < 1/6$,

$$D_H(1/2 - h) = D_G(1/2 - h) - tR_\xi(1/2 - h) = \frac{h^2}{2} + th,$$

which is negative for sufficiently small $h > 0$. Nonnegativity of the masses at $1/6$ and $1/2$ then gives $t \leq 1/4$ which proves (16).

Taking any $t \in (0, 1/4)$ in (16) gives a feasible distribution different from G, supported on $\mathrm{supp}(G)$, and binding wherever G binds. Therefore G is not split-rigid. By Proposition 4, G is not exposed.

To prove extremality of G suppose towards a contradiction that

$$G = \frac{1}{2}H_1 + \frac{1}{2}H_2, \qquad H_1, H_2 \in \mathrm{MPC}(F).$$

Since the H_i are positive measures, every set to which G assigns zero mass must also have zero H_i-mass. Hence

$$\mathrm{supp}(H_i) \subseteq \mathrm{supp}(G).$$

For $x \in K_G$,

$$0 = D_G(x) = \frac{1}{2}D_{H_1}(x) + \frac{1}{2}D_{H_2}(x).$$

Both terms on the right are nonnegative, so

$$D_{H_1}(x) = D_{H_2}(x) = 0.$$

Thus $R_{H_i} = R_F$ on K_G, and (16) gives

$$H_i = G + t_i\xi \qquad \text{for some } t_i \in [0, 1/4].$$

Substituting into the ansatz yields $G = G + \frac{t_1+t_2}{2}\xi$, and since $\xi \neq 0$, this implies that $t_1 + t_2 = 0$. As $t_1, t_2 \geq 0$, this implies that $t_1 = t_2 = 0$, and hence $H_1 = H_2 = G$. Therefore G is extreme.

To show that G is not exposed by any linear functional, define $\rho = -3\delta_{1/6} + 3\delta_{1/3} + \delta_{1/2} - \delta_1$. This signed measure also has zero total mass and zero first moment. Its potential is

$$R_\rho(x) = \begin{cases} 0, & 0 \leq x \leq \dfrac{1}{6}, \\ -3\left(x - \dfrac{1}{6}\right), & \dfrac{1}{6} \leq x \leq \dfrac{1}{3}, \\ -\dfrac{1}{2}, & \dfrac{1}{3} \leq x \leq \dfrac{1}{2}, \\ x - 1, & \dfrac{1}{2} \leq x \leq 1. \end{cases}$$

For $0 < \varepsilon \leq 1/12$, set $J_\varepsilon = G - \varepsilon\xi + \varepsilon^2\rho$. Explicitly,

$$J_\varepsilon = \left(\frac{1}{4} + \varepsilon - 3\varepsilon^2\right)\delta_{1/6} + \left(\frac{1}{4} - 2\varepsilon + 3\varepsilon^2\right)\delta_{1/3} + \left(\frac{1}{4} + \varepsilon + \varepsilon^2\right)\delta_{1/2} + \left(\frac{1}{4} - \varepsilon^2\right)\delta_1.$$

All four coefficients are positive for $0 < \varepsilon \leq 1/12$. Since both ξ and ρ have zero mass and first moment, J_ε is a probability measure with the same mean as F. Its potential gap is

$$D_{J_\varepsilon} = D_G + \varepsilon R_\xi - \varepsilon^2 R_\rho.$$

Using the formulas above,

$$D_{J_\varepsilon}(x) = \begin{cases} \dfrac{x^2}{2}, & 0 \leq x \leq \dfrac{1}{6}, \\ D_G(x) - (\varepsilon - 3\varepsilon^2)\left(x - \dfrac{1}{6}\right), & \dfrac{1}{6} \leq x \leq \dfrac{1}{3}, \\ \dfrac{1}{2}\left(x - \dfrac{1}{2} + \varepsilon\right)^2, & \dfrac{1}{3} \leq x \leq \dfrac{1}{2}, \\ \varepsilon^2(1 - x), & \dfrac{1}{2} \leq x \leq 1. \end{cases}$$

The first, third, and fourth expressions are nonnegative. On $[1/6, 1/3]$, $D_G(x) \geq \frac{1}{96}$. Also, when $0 < \varepsilon \leq 1/12$,

$$0 \leq \varepsilon - 3\varepsilon^2 \leq \frac{1}{16}, \qquad 0 \leq x - \frac{1}{6} \leq \frac{1}{6}.$$

It follows that $D_{J_\varepsilon}(x) \geq \dfrac{1}{96} - \dfrac{1}{16} \cdot \dfrac{1}{6} = 0$. Therefore $J_\varepsilon \in \mathrm{MPC}(F)$ for $\epsilon \in (0, 1/12)$. Moreover, $\frac{J_\varepsilon - G}{\varepsilon} = -\xi + \varepsilon\rho \longrightarrow -\xi$ in total variation.

Let Λ be any real-valued linear functional on the space of finite signed measures, and suppose that G maximizes Λ over $\mathrm{MPC}(F)$.

Since $G + t\xi \in \mathrm{MPC}(F)$ for $0 < t \leq 1/4$, we know that $\Lambda(\xi) \leq 0$. On the other hand, feasibility of J_ε gives

$$\begin{aligned} 0 &\geq \Lambda(J_\varepsilon) - \Lambda(G) \\ &= -\varepsilon\Lambda(\xi) + \varepsilon^2\Lambda(\rho). \end{aligned}$$

Dividing by $\varepsilon > 0$ and letting $\varepsilon \downarrow 0$ yields $\Lambda(\xi) \geq 0$. Consequently, $\Lambda(\xi) = 0$ and it follows that

$$\Lambda(G + t\xi) = \Lambda(G) \qquad \text{for every } 0 \leq t \leq \frac{1}{4}.$$

Thus G cannot be the unique maximizer of Λ.

Note that this argument does not require continuity of Λ or the choice of a topology. In particular, it applies to every setwise-continuous linear functional

$$\Lambda_u(H) = \int u \,\mathrm{d}H, \qquad u \in B_b([0,1]).$$

Hence G is not exposed (under the weak* topology, as defined throughout the paper). As shown, this failure of exposedness is algebraic rather than topological. It survives under the setwise, or total-variation topology; G is not exposed even under the finest locally convex topology.

B.3 Accumulating Point Masses Example

Example 4. The example consists of shrinking one-point posterior components which accumulate at 1. Let

$$x_0 = 0, \qquad x_n = 1 - 2^{-n}, \qquad z_n = x_{n-1} + \frac{x_n - x_{n-1}}{3} = 1 - \frac{5}{3 \cdot 2^n} \quad (n \geq 1).$$

Define the balanced signed measure

$$\eta = \frac{1}{8}\delta_0 + \sum_{n \geq 1} a_n \delta_{z_n}, \qquad a_n = \frac{3(-1)^n}{2^{n+3}},$$

and positive numbers

$$g_n = \begin{cases} |a_n|, & n \text{ odd}, \\ 2^{-n}|a_n|, & n \text{ even}, \end{cases} \qquad g_0 = \frac{29}{40}.$$

Indeed,

$$\sum_{\substack{n \geq 1 \\ n \text{ odd}}} g_n = \frac{1}{4}, \qquad \sum_{\substack{n \geq 1 \\ n \text{ even}}} g_n = \frac{1}{40},$$

so

$$G = g_0\delta_0 + \sum_{n\geq 1} g_n\delta_{z_n}$$

is a probability measure. For each n, spread the mass g_n at z_n, preserving its mean, onto the endpoints of $[x_{n-1}, x_n]$, and set

$$F = g_0\delta_0 + \sum_{n\geq 1} g_n\left(\frac{2}{3}\delta_{x_{n-1}} + \frac{1}{3}\delta_{x_n}\right).$$

Thus F is a countably atomic probability measure and $G \in \mathrm{MPC}(F)$. More explicitly, on each $I_n = (x_{n-1}, x_n)$, the potential gap D_G is a positive triangular function, with its peak at z_n, and it vanishes off the union of these intervals. Hence

$$K_G = \{x_n : n \geq 0\} \cup \{1\}, \qquad \mathrm{supp}(G) = \{0, z_1, z_2, \ldots, 1\}.$$

We next identify a one-sided feasible direction. The cdf of η, hence the slope of R_η, is

$$\frac{1}{8}\left(-\frac{1}{2}\right)^{n-1} \quad \text{on } (x_{n-1}, z_n), \qquad \frac{1}{8}\left(-\frac{1}{2}\right)^{n} \quad \text{on } (z_n, x_n).$$

Since $z_n - x_{n-1} = (x_n - x_{n-1})/3$, these two slopes give zero net change across I_n. Thus $R_\eta = 0$ on K_G; it is positive on the odd intervals and negative on the even ones. The displayed formulas also give $\eta([0,1]) = \int x\,\mathrm{d}\eta(x) = 0$.

For odd n, the two triangular functions agree on I_n:

$$D_G = R_\eta \quad \text{on } I_n.$$

For even n, $D_G > 0 > R_\eta$ on I_n. Moreover,

$$g_n + ta_n = \begin{cases} |a_n|(1-t), & n \text{ odd}, \\ a_n(2^{-n} + t), & n \text{ even}. \end{cases}$$

It follows that

$$G + t\eta \in \mathrm{MPC}(F) \qquad (0 \leq t \leq 1).$$

In particular, G is not split-rigid: for $0 < t < 1$, $G + t\eta \neq G$, its support is contained in $\mathrm{supp}(G)$, and its potential agrees with R_F on K_G.

We now show that this one-sided segment is the whole same-support, same-binding face. Suppose $H \in \mathrm{MPC}(F)$ satisfies $\mathrm{supp}(H) \subseteq \mathrm{supp}(G)$ and $R_H = R_F$ on K_G, and put $\theta = H - G$. On I_n, the function R_θ is affine on either side of its only possible kink z_n, and it vanishes at both endpoints. If $c = \Delta\theta(0)$, induction over the intervals shows that

its slopes are

$$c\left(-\frac{1}{2}\right)^{n-1} \quad \text{and} \quad c\left(-\frac{1}{2}\right)^{n},$$

respectively. Therefore $\theta = t\eta$, where $t = 8c$; the possible atom at the limit point 1 is zero because the slopes tend to zero and θ has total mass zero. Nonnegativity of H on the odd atoms gives $t \le 1$, while on all even atoms it gives $t \ge -2^{-n}$ for every even n, hence $t \ge 0$. We have proved

$$\{H \in \mathrm{MPC}(F) : \mathrm{supp}(H) \subseteq \mathrm{supp}(G),\ R_H = R_F \text{ on } K_G\} = \{G + t\eta : 0 \le t \le 1\}. \quad (17)$$

This makes G extreme. Indeed, if $G = (H_1 + H_2)/2$ with $H_i \in \mathrm{MPC}(F)$, nonnegativity implies $\mathrm{supp}(H_i) \subseteq \mathrm{supp}(G)$. At every $x \in K_G$, the two nonnegative gaps $D_{H_i}(x)$ average to $D_G(x) = 0$, so both vanish. By (17), $H_i = G + t_i\eta$ with $t_i \ge 0$; their average is G only when $t_1 = t_2 = 0$.

It remains to prove the stronger assertion that G is not exposed. For even N, truncate the reverse direction and close it at the binding point x_N:

$$\zeta_N = -\frac{1}{8}\delta_0 - \sum_{n=1}^{N} a_n \delta_{z_n} + \frac{2^{-N}}{8}\delta_{x_N}.$$

Its potential is $-R_\eta$ on $[0, x_N]$ and zero on $[x_N, 1]$. Consequently, for $\varepsilon_N = 2^{-N-1}$,

$$J_N := G + \varepsilon_N \zeta_N \in \mathrm{MPC}(F).$$

To see this directly, on an odd interval the potential gap of J_N is $D_G + \varepsilon_N R_\eta > 0$; on an even interval I_n, $n \le N$, it is $D_G - \varepsilon_N |R_\eta| \ge 0$ because $D_G/|R_\eta| = g_n/a_n = 2^{-n} \ge 2^{-N} > \varepsilon_N$; after x_N the potential gap is unchanged. The same inequalities show nonnegativity of the atom weights, and the added atom at x_N is positive. Also

$$\zeta_N \longrightarrow -\eta \quad \text{in total variation.}$$

Let Λ be a linear functional on the finite signed Borel measures on $[0, 1]$ that is continuous in total variation and attains its maximum over $\mathrm{MPC}(F)$ at G. This covers $\Lambda(H) = \int u \, \mathrm{d}H$ for every bounded measurable u, and a fortiori every weak*- or setwise-continuous linear functional. Feasibility of $G + t\eta$ for $t > 0$ gives $\Lambda(\eta) \le 0$. Feasibility of J_N gives $\Lambda(\zeta_N) \le 0$; since $\zeta_N \to -\eta$ in total variation, passing to the limit gives $-\Lambda(\eta) \le 0$. Hence $\Lambda(\eta) = 0$, so every $G + t\eta$, $0 < t \le 1$, is another maximizer. No linear functional that is continuous in total variation exposes G; in particular, G is not exposed under the weak*, setwise, or total-variation topology.

Exposure under the finest locally convex topology

The failure of exposedness just shown is topological, not algebraic. We now construct a linear functional that fails to be continuous in total-variation and has G as its unique maximizer over $\mathrm{MPC}(F)$. Since every linear functional is continuous under the finest locally convex topology τ_{flc}, G is exposed under τ_{flc}. This contrasts with Example 2, where no linear functional at all exposes G.

Throughout, for $n \geq 1$ let

$$r_n(H) = \frac{H(\{z_n\})}{|a_n|}$$

denote the mass of H at z_n measured in units of the atom of η there. By definition of g_n and a_n,

$$r_n(G) = \begin{cases} 1, & n \text{ odd}, \\ 2^{-n}, & n \text{ even}, \end{cases} \qquad r_n(G + t\eta) = 1 - t \quad \text{for odd } n.$$

Fix $H \in \mathrm{MPC}(F)$ and $N \geq 0$. Since H and F are probability measures with the same mean e_F, the decomposition $(t - x_N)_+ = (t - x_N) + (x_N - t)_+$ and Lemma 1 give

$$\int (t - x_N)_+ \, \mathrm{d}H(t) = e_F - x_N + R_H(x_N),$$

and likewise for F. The feasibility constraint $R_H(x_N) \leq R_F(x_N)$ is therefore equivalent to

$$\int (t - x_N)_+ \, \mathrm{d}H(t) \leq \int (t - x_N)_+ \, \mathrm{d}F(t). \tag{18}$$

We evaluate both integrals. The integrand vanishes on $[0, x_N]$, which contains the atom of F at 0 and the blocks $n \leq N$. For $n > N$ both x_{n-1} and x_n lie in $[x_N, 1]$, where the integrand equals $t - x_N$, and the prior mass g_n on block n has barycenter z_n. Hence

$$\int (t - x_N)_+ \, \mathrm{d}F(t) = \sum_{n>N} g_n(z_n - x_N). \tag{19}$$

Since the integrand is nonnegative and equals $z_n - x_N$ at z_n for $n > N$, keeping only the mass of H at the points z_n, $n > N$, gives

$$\int (t - x_N)_+ \, \mathrm{d}H(t) \geq \sum_{n>N} H(\{z_n\})(z_n - x_N). \tag{20}$$

Combining (18)–(20),

$$\sum_{n>N} \Big(H(\{z_n\}) - g_n\Big)(z_n - x_N) \leq 0. \tag{21}$$

We rewrite (21) in terms of the ratios r_n. For $n = N + k$ with $k \geq 1$,

$$z_n - x_N = 2^{-N} w_k, \qquad w_k := 1 - \tfrac{5}{3} 2^{-k} \in [\tfrac{1}{6}, 1),$$

and

$$H(\{z_n\}) - g_n = |a_n|\Big(r_n(H) - r_n(G)\Big), \qquad |a_n| = 3 \cdot 2^{-N-3}\, 2^{-k}.$$

Substituting into (21) and dividing by the positive constant $3 \cdot 2^{-2N-3}$ yields

$$\sum_{k\geq 1} 2^{-k} w_k \Big(r_{N+k}(H) - r_{N+k}(G)\Big) \leq 0 \qquad \text{for every } N \geq 0. \tag{22}$$

Thus (22) is nothing but the binding constraint $D_H(x_N) \geq 0$ at the point $x_N \in K_G$, rescaled and weakened by discarding the mass of H off the points z_n.

In (22), keep the term $k = 1$ on the left and move the terms $k \geq 2$ to the right. Dropping $r_{N+k}(H) \geq 0$ and using $w_k < 1$ and $r_n(G) \leq 1$,

$$\tfrac{1}{2} w_1 \Big(r_{N+1}(H) - r_{N+1}(G)\Big) \leq \sum_{k\geq 2} 2^{-k} w_k\, r_{N+k}(G) \leq \sum_{k\geq 2} 2^{-k} = \tfrac{1}{2}.$$

Since $w_1 = \frac{1}{6}$, this gives $r_{N+1}(H) \leq r_{N+1}(G) + 6 \leq 7$. As $N \geq 0$ was arbitrary,

$$0 \leq r_n(H) \leq 7 \qquad \text{for all } n \geq 1 \text{ and all } H \in \mathrm{MPC}(F). \tag{23}$$

Let $s_j(H) = r_{2j+1}(H)$ for $j \geq 0$ be the sequence of ratios at the odd atoms. Let LIM be a Banach limit on ℓ^∞, that is, a linear functional satisfying $\inf_j b_j \leq \mathrm{LIM}(b) \leq \sup_j b_j$ and $\mathrm{LIM}(Sb) = \mathrm{LIM}(b)$, where $(Sb)_j = b_{j+1}$ is the shift. In particular, LIM is positive, has norm one, and equals the limit on convergent sequences. On the linear subspace

$$\mathcal{V} = \Big\{\theta \in \mathcal{M} : (s_j(\theta))_{j\geq 0} \in \ell^\infty\Big\}$$

define $\Lambda_{\mathrm{sing}}(\theta) = \mathrm{LIM}_j s_j(\theta)$, and extend Λ_{sing} linearly from $\mathcal{V}$ to all of $\mathcal{M}$. By (23), $\mathrm{MPC}(F) \subseteq \mathcal{V}$. Since $s_j(G) = 1$ and $s_j(\eta) = a_{2j+1}/|a_{2j+1}| = -1$ for every j,

$$\Lambda_{\mathrm{sing}}(G) = 1, \qquad \Lambda_{\mathrm{sing}}(\eta) = -1, \qquad \Lambda_{\mathrm{sing}}(G + t\eta) = 1 - t.$$

We claim that

$$\Lambda_{\mathrm{sing}}(H) \leq 1 \qquad \text{for every } H \in \mathrm{MPC}(F). \tag{24}$$

Fix $H \in \mathrm{MPC}(F)$ and $m \geq 0$, and apply (22) with $N = 2m$. Split the sum according to the parity of k.

- For even k, the index $2m + k$ is even, so $r_{2m+k}(G) = 2^{-2m-k}$, while $r_{2m+k}(H) \geq 0$ and $w_k < 1$. Each such term is therefore at least $-2^{-2m} 4^{-k}$, and summing over even $k \geq 2$,

$$\sum_{k \text{ even}} 2^{-k} w_k \Big(r_{2m+k}(H) - r_{2m+k}(G)\Big) \geq -2^{-2m} \sum_{j\geq 1} 16^{-j} = -\frac{2^{-2m}}{15}.$$

- For odd $k = 2i-1$, $i \geq 1$, the index is $2m+2i-1 = 2(m+i-1)+1$, so $r_{2m+k}(G) = 1$ and $r_{2m+k}(H) = s_{m+i-1}(H)$. With $v_i := 2^{-(2i-1)} w_{2i-1} > 0$,

$$\sum_{k \text{ odd}} 2^{-k} w_k \Big(r_{2m+k}(H) - r_{2m+k}(G) \Big) = \sum_{i \geq 1} v_i \Big(s_{m+i-1}(H) - 1 \Big).$$

Inequality (22) says that the odd sum plus the even sum is at most 0. Hence the odd sum is at most minus the even sum, that is,

$$\sum_{i \geq 1} v_i \Big(s_{m+i-1}(H) - 1 \Big) \leq \frac{2^{-2m}}{15} \qquad \text{for every } m \geq 0. \tag{25}$$

Both sides of (25) are bounded sequences in m, and we apply LIM to each. The sequence $m \mapsto 2^{-2m}/15$ converges to 0, so its Banach limit is 0. On the left, the i-th summand is the sequence $m \mapsto s_{m+i-1}(H) - 1$, which is the $(i-1)$-fold shift of $s(H) - \mathbf{1}$ and has sup norm at most 6 by (23). Since $\sum_i v_i < \infty$, the series converges in ℓ^∞, so by norm continuity, linearity, and shift invariance of LIM,

$$\mathrm{LIM}_m \sum_{i \geq 1} v_i \Big(s_{m+i-1}(H) - 1 \Big) = \sum_{i \geq 1} v_i \Big(\mathrm{LIM}_j \, s_j(H) - 1 \Big) = \Big(\sum_{i \geq 1} v_i \Big) \Big(\Lambda_{\mathrm{sing}}(H) - 1 \Big).$$

Positivity of LIM applied to (25) therefore gives $\Big(\sum_i v_i \Big) \Big(\Lambda_{\mathrm{sing}}(H) - 1 \Big) \leq 0$, which is (24).

Let u be the Lipschitz objective constructed in the proof of Proposition 3 for the present G; the construction uses only K_G and $\mathrm{supp}(G)$, not split rigidity. Put $\Lambda_u(H) = \int u \, \mathrm{d}H$. By (7), $\Lambda_u(H) \leq \Lambda_u(G)$ for all $H \in \mathrm{MPC}(F)$, with equality if and only if $\mathrm{supp}(H) \subseteq \mathrm{supp}(G)$ and $R_H = R_F$ on K_G. By (17), this holds if and only if $H = G + t\eta$ for some $t \in [0,1]$. In particular $\Lambda_u(\eta) = 0$.

Set $\Lambda = \Lambda_u + \Lambda_{\mathrm{sing}}$. For $H \in \mathrm{MPC}(F)$, (24) gives

$$\Lambda(H) = \Lambda_u(H) + \Lambda_{\mathrm{sing}}(H) \leq \Lambda_u(G) + 1 = \Lambda(G).$$

If equality holds, then $\Lambda_u(H) = \Lambda_u(G)$, so $H = G + t\eta$ with $t \in [0,1]$, and $\Lambda_{\mathrm{sing}}(H) = 1 - t = 1$, so $t = 0$. Thus G is the unique maximizer of Λ over $\mathrm{MPC}(F)$, and G is exposed under τ_{flc}.

As it must be, Λ is not continuous in total variation: $\Lambda_{\mathrm{sing}}(\zeta_N) = 0$ since ζ_N has only finitely many odd atoms, so $\Lambda(\zeta_N) = \Lambda_u(\zeta_N) \to -\Lambda_u(\eta) = 0$, whereas $\Lambda(-\eta) = -\Lambda_u(\eta) - \Lambda_{\mathrm{sing}}(\eta) = 1$, although $\zeta_N \to -\eta$ in total variation.

B.4 When are all extreme points exposed?

Proposition 8 (Isolated atoms and finitely many pooling intervals)**.** *Suppose every atom x of F is isolated in* $\mathrm{supp}(F)$*: there exists a neighborhood of x that contains no other*

F-mass. If $G \in \mathrm{MPC}(F)$ is extreme and has finitely many pooling intervals, then G is split-rigid and hence exposed.

Proof. Let $I_1, \dots, I_N$ be the pooling intervals, and let $H \in \mathrm{MPC}(F)$ satisfy

$$\mathrm{supp}(H) \subseteq \mathrm{supp}(G), \qquad R_H = R_F \quad \text{on } K_G.$$

We prove that $H = G$.

Since G is extreme, Proposition 2 implies that G is bi-pooling. Hence, on each I_j, both G and H are carried by the same set of at most two atoms. On $\operatorname{int} K_G$, the equality

$$R_H = R_F = R_G$$

on a neighborhood implies that $H = G$.

Consider an endpoint x of a pooling interval. If $\Delta F(x) = 0$, the endpoint-slope inequalities applied to H and G give

$$F(x-) = H(x-) = H(x) = G(x-) = G(x) = F(x),$$

with the evident one-sided interpretation at 0 and 1. Thus $H - G$ has no atom at x. If $\Delta F(x) > 0$, then x is isolated in $\mathrm{supp}(F)$. Since the adjacent posterior components have finite support, x is also isolated in $\mathrm{supp}(G)$; therefore support inclusion allows H to charge x only if $\Delta G(x) > 0$. It follows that, for some $C \geq 1$,

$$H \leq CG.$$

We next enlarge C, if necessary, so that

$$|R_H - R_G| \leq C D_G \qquad \text{on } [0,1].$$

The inequality is immediate on K_G, where $R_H = R_G = R_F$. On a pooling interval $I = (a, b)$, the function $R_H - R_G$ is piecewise affine with finitely many breakpoints and vanishes at a and b. At a nonatomic endpoint it vanishes on a one-sided neighborhood of the endpoint. At an isolated atomic endpoint, for all sufficiently small $h > 0$,

$$D_G(a + h) = \lambda_I^a \, h, \qquad D_G(b - h) = \lambda_I^b \, h,$$

respectively. The shares are strictly positive, since otherwise D_G would vanish on a subinterval of I. Since $R_H - R_G$ is also linear there, the required ratio is bounded near the endpoints. Away from the endpoints, D_G has a positive minimum. The bound therefore holds on each pooling interval, and a common C exists because there are only finitely many of them.

Choose $0 < \varepsilon < 1/C$. Then

$$(1+\varepsilon)G - \varepsilon H \geq \Big(1 - \varepsilon(C-1)\Big)G \geq 0.$$

This measure has total mass one and the same first moment as F. Moreover,

$$\begin{aligned} R_{(1+\varepsilon)G-\varepsilon H} &= R_G + \varepsilon(R_G - R_H) \\ &\leq R_G + \varepsilon|R_G - R_H| \\ &\leq R_G + \varepsilon C D_G \\ &\leq R_G + D_G = R_F. \end{aligned}$$

Thus $(1+\varepsilon)G - \varepsilon H \in \mathrm{MPC}(F)$. Since

$$G = \frac{\varepsilon}{1+\varepsilon}H + \frac{1}{1+\varepsilon}\Big((1+\varepsilon)G - \varepsilon H\Big),$$

extremality of G forces $H = G$. Hence G is split-rigid, and Proposition 3 gives exposedness. □

Corollary 5 (Finitely supported priors)**.** *If* $|\operatorname{supp}(F)| = n < \infty$*, then the split-rigid, exposed, and extreme points of* $\mathrm{MPC}(F)$ *coincide. Every feasible distribution has at most* $n-1$ *pooling intervals, and this bound is attained. The inclusion of these classes in the bi-pooling laws can nevertheless be strict (Example 1).*

Proof. Write $\operatorname{supp}(F) = \{x_1 < \cdots < x_n\}$ and let I be a pooling interval of some $G \in \mathrm{MPC}(F)$. By (2) and Lemma 5, the prior component F_I is carried by $\overline{I} \cap \operatorname{supp}(F)$. It is not a point mass: the only mean-preserving contraction of a point mass is itself, so $F_I = m\delta_x$ would give $G_I = F_I$ and, by (11), $D_G = 0$ on I. Hence $\overline{I}$ contains two support points, and the open interval I contains the gap (x_j, x_{j+1}) between some consecutive pair. Disjoint pooling intervals contain distinct gaps, and there are $n-1$ gaps. The bound is attained by pooling every gap: assign the whole atom at x_1 and at x_n, and half of every other atom, to each adjacent gap, and contract the mass assigned to (x_j, x_{j+1}) to its barycenter. The potential gap is then a positive tent on every (x_j, x_{j+1}) and vanishes at every x_j. All atoms are isolated, so Proposition 8 and Theorem 2 give the coincidence of the three classes. The reverse inclusions follow since all exposed points are extreme. □

B.5 Non-Exposed Extreme Points can be Uniquely Optimal under USC objectives

Example 5 (A uniquely optimal distribution that is not split-rigid)**.** Unique optimality for an upper semicontinuous objective does not imply split rigidity. To see this, let

$$a = \frac{3}{16}, b = \frac{5}{16}, c = \frac{1}{2}, \qquad F = \tfrac{1}{2}\,\mathsf{U}[0,c] + \tfrac{1}{2}\delta_c, \qquad G = \frac{1}{4}\delta_a + \frac{1}{4}\delta_b + \frac{1}{2}\delta_c.$$

Both distributions have mean $3/8$. A direct calculation gives $D_G(x) = x^2/2$ on $[0,a]$, $D_G(x) = (x-1/4)^2/2 + 1/64$ on $[a,b]$, and $D_G(x) = (c-x)^2/2$ on $[b,c]$; moreover, $D_G = 0$ on $[c,1]$. Thus $G \in \mathrm{MPC}(F)$ and $K_G = \{0\} \cup [c,1]$.

For $0 < \varepsilon < 1/6$, define

$$H_\varepsilon = \left(\frac{1}{4} - \frac{3\varepsilon}{2}\right)\delta_a + \left(\frac{1}{4} + \frac{5\varepsilon}{2}\right)\delta_b + \left(\frac{1}{2} - \varepsilon\right)\delta_c.$$

The coefficients are positive, and H_ε has the same mass and mean as G. Its potential differs from that of G by zero below a and above c, by $-3\varepsilon(x-a)/2$ on $[a,b]$, and by $\varepsilon(x-c)$ on $[b,c]$. Hence $R_{H_\varepsilon} \le R_G \le R_F$, so $H_\varepsilon \in \mathrm{MPC}(F)$. In addition, $\mathrm{supp}(H_\varepsilon) = \mathrm{supp}(G)$ and $R_{H_\varepsilon} = R_F$ on K_G. Since $H_\varepsilon \neq G$, the distribution G is not split-rigid.

Nevertheless, G is uniquely optimal for a bounded upper semicontinuous objective. Let

$$u(x) = \begin{cases} 1, & x = c, \\ 0, & x \in \{a,b\}, \\ -1, & x \notin \{a,b,c\}. \end{cases}$$

The function u is upper semicontinuous: its upper level sets are closed.

Now take any $H \in \mathrm{MPC}(F)$. Every feasible H is supported on $[0,c]$. By (1), $D'_H(c-) = F(c-) - H(c-) = \Delta H(c) - 1/2$. Since $D_H \ge 0$ and $D_H(c) = 0$, this left derivative is nonpositive, and therefore $\Delta H(c) \le 1/2$. Consequently, $\int u\,\mathrm{d}H = \Delta H(c) - H([0,1] \setminus \{a,b,c\}) \le 1/2 = \int u\,\mathrm{d}G$.

Equality requires both $\Delta H(c) = 1/2$ and $H([0,1] \setminus \{a,b,c\}) = 0$. Thus H is supported on $\{a,b,c\}$ and puts mass $1/2$ at c. Total mass and the mean then force the remaining masses to be $1/4$ at a and $1/4$ at b. Hence $H = G$. Thus G is the unique maximizer of a bounded upper semicontinuous objective, although it is not split-rigid.