Distributions of posterior quantiles via matching

ANTON KOLOTILIN
School of Economics, UNSW Business School

ALEXANDER WOLITZKY
Department of Economics, MIT

We offer a simple analysis of the problem of choosing a statistical experiment to optimize the induced distribution of posterior medians, or more generally q-quantiles for any $q \in (0,1)$. We show that a single experiment—the q-*quantile matching experiment*—implements all implementable distributions of posterior q-quantiles, with different distributions spanned by different selections from the sets of posterior q-quantiles. A dense subset of implementable distributions of posterior q-quantiles can be uniquely implemented by perturbing the q-quantile matching experiment. A linear functional is optimized over distributions of posterior q-quantiles by taking the optimal selection from each set of posterior q-quantiles. The q-quantile matching experiment is the only experiment that simultaneously implements all implementable distributions of posterior q-quantiles.

KEYWORDS. quantiles, statistical experiments, median matching, overconfidence, gerrymandering, persuasion.

JEL CLASSIFICATION. C61, D72, D82.

1. INTRODUCTION

Several problems of recent economic interest amount to characterizing the set of distributions of posterior quantiles that can be induced by some statistical experiment, or to finding a distribution in this set that maximizes some objective. These problems include *apparent overconfidence* (Benoît and Dubra, 2011)—e.g., what distributions of medians of individuals' beliefs about their own abilities are consistent with Bayesian updating?; *partisan gerrymandering* (Friedman and Holden 2008; Kolotilin and Wolitzky 2020b)—e.g., what is the highest distribution of district median voters attained by any districting plan?; and *quantile persuasion* (Kolotilin and Wolitzky, 2020a)—e.g., what experiment maximizes the expected action of a receiver who minimizes the expected absolute deviation of her action from the unknown state of the world?[1]

Our problem is as follows. There is a real-valued state θ. A statistical experiment induces a distribution over posteriors μ. For any $q \in (0,1)$, each posterior μ has at least

Anton Kolotilin: akolotilin@gmail.com
Alexander Wolitzky: alexander.wolitzky@gmail.com

[1] Yang and Zentefis (2024) explore these and other applications. Kolotilin and Wolitzky (2020b) consider a more general gerrymandering model, which reduces to optimizing the distribution of posterior quantiles in a special case. Kolotilin and Wolitzky (2020a, Proposition 2') introduce quantile persuasion as a special case of a more general persuasion model, which is further developed in Kolotilin et al. (2024).

one q-quantile. In general, a posterior can have multiple q-quantiles due to gaps in the support of μ: for example, if μ puts equal weight on two states $\theta < \theta'$, then the set of medians of μ is the entire interval $[\theta, \theta']$. An experiment, together with a selection rule to break ties for posteriors with multiple q-quantiles, induces a distribution of posterior q-quantiles. A distribution of posterior q-quantiles is *implementable* if it is induced by some experiment and selection rule; it is *uniquely implementable* if it is induced by an experiment that almost always induces posteriors with unique q-quantiles. We ask what distributions of posterior q-quantiles are implementable or uniquely implementable; how to implement them; and how to optimize a linear functional over distributions of posterior q-quantiles.

We provide a simple solution to this problem. For any $q \in (0,1)$, there is a single experiment—the *q-quantile matching experiment*—that simultaneously implements all implementable distributions of posterior q-quantiles, with different distributions spanned by different selection rules. For example, if the state is uniformly distributed on $[0,1]$ and the relevant quantile is the median, the q-quantile matching experiment is the *median matching experiment* that, whenever the true state is $\theta \in [0, 1/2]$, reveals only that the state is either θ or $1/2+\theta$ (and, hence, whenever the true state is $\theta \in (1/2, 1]$, reveals only that the state is either θ or $\theta - 1/2$).[2] In general, the q-quantile matching experiment pools pairs of states across a q-quantile of the prior in a positively assortative manner, with weight q on the lower state in each pair.

To see why the q-quantile matching experiment implements all implementable distributions of posterior q-quantiles, consider again the median matching experiment with a uniform state. When the experiment reveals that the state is θ or $1/2+\theta$ with equal probability, every value $x \in [\theta, 1/2+\theta]$ is a posterior median. The median matching experiment thus simultaneously implements (i) the distribution $\underline{H}(x) = \max\{0, 2x-1\}$, (ii) the distribution $\overline{H}(x) = \min\{2x, 1\}$, and (iii) every distribution H satisfying $\underline{H} \leq H \leq \overline{H}$. Conversely, simple Markov-type inequalities imply that every implementable distribution is bounded by $\underline{H}$ and $\overline{H}$. Moreover, the set of uniquely implementable distributions of posterior quantiles is essentially the same: any desired selection from each set of q-quantiles induced by the q-quantile matching experiment can be uniquely selected by mixing each posterior under the q-quantile matching with the degenerate distribution on the desired selection with probabilities $1-\varepsilon$ and ε, respectively. Finally, optimizing a linear functional over distributions of posterior quantiles simply requires taking the optimal selection from each set of q-quantiles induced by the q-quantile matching experiment. See Figure 1 for an illustration of our results.[3]

We also show that the q-quantile matching experiment is the *unique* experiment that simultaneously implements all implementable distributions of posterior q-quantiles. To see why, consider again a uniform state, and compare the median matching experiment with the *negative assortative matching* experiment that, whenever the true state

[2]To our knowledge, the median matching experiment first appears in Kolotilin and Wolitzky (2020a, p. 29). It is closely related to the median one-to-one matching introduced by Kremer and Maskin (1996) and further studied by Legros and Newman (2002)—the title of the present paper acknowledges this connection.

[3]Similar figures in the literature include Figure 1 of Owen and Grofman (1988), Figure 2 of Kamenica and Gentzkow (2011), and Figure 3 of Yang and Zentefis (2024).

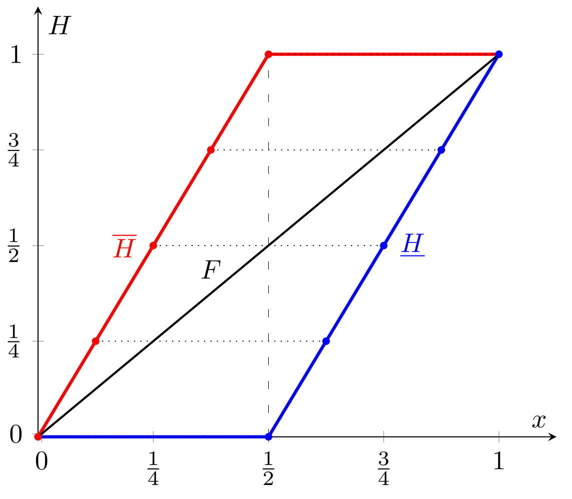

FIGURE 1. Implementable Distributions of Posterior Medians

FIGURE 1. *Notes:* When the prior F is uniform on $[0,1]$, $\overline{H}$ and $\underline{H}$ are the lowest and highest implementable distributions of posterior medians. A distribution H is implementable iff $\underline{H} \leq H \leq \overline{H}$. Optimizing a linear functional over distributions of posterior medians requires taking the optimal selection from each horizontal dotted line. For example, the blue (red) dots are the optimal selections for an increasing (decreasing) objective function.

is $\theta \in [0,1]$, reveals only that the state is either θ or $1-\theta$. The negative assortative matching experiment simultaneously implements the lowest and highest distributions of posterior medians, $\overline{H}$ and $\underline{H}$, but it does not implement all intermediate distributions, such as the distribution $H_{1/2}$ given by $H_{1/2}(x) = \overline{H}(x)$ for $x < 1/4$, $H_{1/2}(x) = 1/2$ for $x \in [1/4, 3/4)$, and $H_{1/2}(x) = \underline{H}(x)$ for $x \geq 3/4$. Indeed, the negative assortative matching experiment induces posteriors with medians between $1/4$ and $3/4$ when the true state lies between $1/4$ and $3/4$, while $H_{1/2}$ assigns probability 0 to these medians.

The current paper is closely related to Benoît and Dubra (2011) and Yang and Zentefis (2024). Both of these papers establish results that are very similar to our Theorem 1 (albeit Benoît and Dubra do so for discrete experiments with finitely many induced posteriors). Our main contribution is introducing the q-quantile matching experiment, which yields a much simpler proof of Theorem 1, as well as new results (Theorems 2 and 3).

2. IMPLEMENTABLE DISTRIBUTIONS OF POSTERIOR QUANTILES

This section shows that the q-quantile matching experiment implements all implementable distributions of posterior q-quantiles.

Let $\Theta = [\underline{\theta}, \overline{\theta}] \subset \mathbb{R}$, with $\underline{\theta} < \overline{\theta}$, be a compact state space; let $C(\Theta)$ be the set of continuous functions on Θ; let $\Delta(\Theta)$ be the set of cumulative distribution functions on Θ, endowed with the weak* topology; and let $\Delta(\Delta(\Theta))$ be the set of probability measures on $\Delta(\Theta)$. Recall that $G \in \Delta(\Theta)$ is a non-decreasing, right-continuous function satisfying

$G(\underline{\theta}) \geq 0$ and $G(\overline{\theta}) = 1$. Let δ_x, with $x \in \Theta$, denote the degenerate distribution at x, so that $\delta_x(\theta) = \mathbf{1}\{\theta \geq x\}$.

Fix a prior distribution $F \in \Delta(\Theta)$ and a quantile of interest $q \in (0,1)$. Following Kamenica and Gentzkow (2011), define an *experiment* as a distribution $\tau \in \Delta(\Delta(\Theta))$ of posterior distributions $G \in \Delta(\Theta)$ such that $\int G \mathrm{d}\tau(G) = F$. For each posterior G, define the set of *q-quantiles* of G as

$$X(G) = \{x \in \Theta : G(x^-) \leq q \leq G(x)\},$$

where $G(x^-)$ denotes the left limit $\lim_{\theta \uparrow x} G(\theta)$, with the convention $G(\underline{\theta}^-) = 0$. In addition, for each G, define its generalized inverse G^{-1} as

$$G^{-1}(p) = \inf\{\theta \in \Theta : G(\theta) \geq p\}, \quad \text{for all } p \in [0,1].$$

That is, $G^{-1}(p)$ is the smallest p-quantile of G.

To define the q-quantile matching experiment, let ω be uniformly distributed on $[0,1]$, and for each $\omega \in [0,q]$, let $G = G_\omega$ be the distribution that assigns probability q to $F^{-1}(\omega)$ and assigns probability $1-q$ to $F^{-1}(q + (1-q)\omega/q)$. The *$q$-quantile matching experiment* is defined as an experiment $\tau^\star$ such that for $\tau^\star$-almost all G, there exists $\omega \in [0,q]$ such that $G = G_\omega$.[4] Formally, $\tau^\star$ is defined by

$$\tau^\star(M) = \int_0^q \mathbf{1}\{q\delta_{F^{-1}(\omega)} + (1-q)\delta_{F^{-1}(q + \frac{1-q}{q}\omega)} \in M\} \frac{\mathrm{d}\omega}{q}, \quad \text{for all } M \subset \Delta(\Theta).$$

While all of our results hold for general F and q, for simplicity we will provide intuition only for the uniform-median case where F is uniform on $[0,1]$ and $q = 1/2$.

A distribution H of q-quantiles is *implemented* by an experiment τ if there exists a (measurable) selection χ from the correspondence X such that the distribution of $\chi(G)$ induced by τ is H. A distribution H of q-quantiles is *uniquely implemented* by an experiment τ if H is implemented by τ and $X(G)$ is a singleton for τ-almost all G. Let $\mathcal{H}$ and $\mathcal{H}^\star$ be the sets of implementable and uniquely implementable distributions of q-quantiles.

The following theorem characterizes $\mathcal{H}$ and $\mathcal{H}^\star$.

THEOREM 1. *The following hold:*

(i) $\mathcal{H} = \{H \in \Delta(\Theta) : \underline{H} \leq H \leq \overline{H}\}$, *where* $\underline{H}(x) = \max\{0, (F(x)-q)/(1-q)\}$ *and* $\overline{H}(x) = \min\{F(x)/q, 1\}$ *for all* $x \in \Theta$.

(ii) *Every* $H \in \mathcal{H}$ *is implemented by* $\tau^\star$.

(iii) *If* F *has a positive density on* Θ *then* $\mathcal{H}$ *is the closure of* $\mathcal{H}^\star$. *In particular, for any objective function* $V \in C(\Theta)$, *we have*

$$\sup_{H \in \mathcal{H}^\star} \int_\Theta V(x)\mathrm{d}H(x) = \max_{H \in \mathcal{H}} \int_\Theta V(x)\mathrm{d}H(x). \tag{1}$$

[4] For example, when F is atomless, we can let $\omega = F(\theta)$, so that the q-quantile matching experiment induces posteriors that assign probability q to θ and assign probability $1-q$ to $F^{-1}(q + (1-q)F(\theta)/q)$, for $\theta \in [0, F^{-1}(q)]$.

Figure 1 illustrates the set $\mathcal{H}$. The intuition for Theorem 1 is straightforward. First, by simple Markov-type inequalities, any implementable H must satisfy $\underline{H} \le H \le \overline{H}$. For example, if the posterior median is less than x with probability p, then θ must be less than x with probability at least $p/2$. When $F(x) = x$, this implies that $p \le 2x$, so the probability that the posterior median is less than x is at most $\min\{2x, 1\} = \overline{H}(x)$.[5]

Conversely, to see that any H satisfying $\underline{H} \le H \le \overline{H}$ is implementable, consider the median matching experiment $\tau^\star$ that induces only posteriors G_θ that assign equal probability to some $\theta \in [0, 1/2]$ and to $1/2 + \theta$. The set of medians of such a posterior is $X(G_\theta) = [\theta, 1/2 + \theta]$. At the same time, $H \le \overline{H}$ implies that $H^{-1}(2\theta) \ge \theta$, and $H \ge \underline{H}$ implies that $H^{-1}(2\theta) \le 1/2 + \theta$, so we have $H^{-1}(2\theta) \in [\theta, 1/2 + \theta]$. Thus, $\chi(G_\theta) = H^{-1}(2\theta)$ is a selection from $X(G_\theta)$. Finally, the distribution of $\chi(G_\theta)$ induced by $\tau^\star$ is H, because the states that induce medians below x under $\tau^\star$ with selection $\chi(G_\theta)$ are precisely those in $[0, H(x)/2]$ and $[1/2, 1/2 + H(x)/2]$, and the measure of these states is $H(x)$.

As to unique implementation, for any $e \in (0, 1]$ and any implementable and absolutely continuous distribution H with density h, we explicitly construct a modification of the median matching experiment $\tau_e^\star$ that uniquely implements the distribution $(1 - e)H + eF$ of medians, by making every posterior G_θ a convex combination of the median matching distribution $(\delta_\theta + \delta_{1/2+\theta})/2$ and the degenerate distribution $\delta_{H^{-1}(2\theta)}$ at the unique median $H^{-1}(2\theta) \in [\theta, 1/2 + \theta]$. Intuitively, for each $\theta \in [0, 1/2]$, $\tau_e^\star$ induces posteriors G_θ and $G_{H(\theta)/2}$ with probabilities $1 - e$ and e; similarly, for each $\theta \in (1/2, 1]$, $\tau_e^\star$ induces posteriors $G_{\theta-1/2}$ and $G_{H(\theta)/2}$ with probabilities $1 - e$ and e. Then posterior medians in $[x, x + \mathrm{d}x]$ are induced at $\theta \in [H(x)/2, H(x + \mathrm{d}x)/2]$ with probability $1 - e$, at $\theta \in [1/2 + H(x)/2, 1/2 + H(x + \mathrm{d}x)/2]$ with probability $1 - e$, and at $\theta \in [x, x + \mathrm{d}x]$ with probability e. Since $H(x + \mathrm{d}x) = H(x) + h(x)\mathrm{d}x$, the density of the posterior median x multiplied by the posterior at x is equal to $(1 - e)h(x)(\delta_{H(x)/2} + \delta_{1/2+H(x)/2})/2 + e\delta_x$, as required. To complete the proof of Theorem 1, we provide a simple argument showing that any distribution in $\mathcal{H}$ can be approximated by uniquely implementable distributions $(1 - e)H + eF$.[6]

The literature contains several close antecedents of Theorem 1. Friedman and Holden (2008) study partisan gerrymandering with a finite number of legislative districts. Benoît and Dubra (2011) study testing for overconfidence in a self-ranking experiment with a finite number of bins. In our notation, Friedman and Holden and Benoît and Dubra consider discrete experiments with finitely many induced posteriors. Friedman and Holden show that a discrete version of $\underline{H}$ is the highest implementable distribution of posterior medians. Benoît and Dubra show that the set of uniquely implementable distributions of posterior medians is a discrete version of the set $\{H \in \Delta(\Theta) : \underline{H} < H < \overline{H}\}$. In a general setting with possibly infinitely many induced

[5]This argument is closely related to Kamenica and Gentzkow's 2011 "prosecutor-judge" example. As in their example, the key observation is that if the prior probability of an event (e.g., the event that $\theta \le x$) is x, then the maximum probability that the posterior probability of this event is at least $1/2$ is $\min\{2x, 1\}$.

[6]A complete characterization of the set $\mathcal{H}^\star$ remains an open problem. Two observations are that $\mathcal{H}^\star$ is a proper subset of $\mathcal{H}$ (as $\underline{H}$ and $\overline{H}$ do not belong to $\mathcal{H}^\star$), and that not all uniquely implementable distributions can be implemented by our modification of q-quantile matching. For example, $H = \delta_{1/2}$ is uniquely implemented by complete pooling but not by our modification of median matching.

posteriors in the contexts of quantile persuasion and partisan gerrymandering, respectively, Kolotilin and Wolitzky (2020a) and Kolotilin and Wolitzky (2020b) show that $\underline{H}$ is the highest implementable distribution of posterior medians. Finally, in a general setting, Yang and Zentefis (2024) show that the set of implementable distributions of posterior medians is $\{H \in \Delta(\Theta) : \underline{H} \leq H \leq \overline{H}\}$, and also construct a dense subset of distributions that are uniquely implementable.[7] Relative to Benoît and Dubra and Yang and Zentefis, Theorem 1 shows that the q-quantile matching experiment implements every $H \in \mathcal{H}$, and also yields a much simpler proof.

Farther afield, Blackwell (1953), Strassen (1965), and Kolotilin (2018) characterize implementable distributions of posterior *means*. An interesting open question is whether a useful analogue of Theorem 1 (for medians) and Strassen's theorem (for means) exists for intermediate statistics that interpolate between the median and the mean.[8]

3. Optimal distributions of posterior quantiles

This section uses the q-quantile matching experiment to characterize the distributions of posterior q-quantiles that maximize a continuous linear functional.

THEOREM 2. *Let $V \in C(\Theta)$. Then H (uniquely) maximizes $\int V(x)\mathrm{d}H(x)$ on $\mathcal{H}$ iff $H^{-1}(p)$ (uniquely) maximizes V on $[\overline{H}^{-1}(p), \underline{H}^{-1}(p)]$ for (almost) all $p \in [0,1]$. Consequently, the value of the maximization problem is*

$$\max_{H\in\mathcal{H}} \int_{\Theta} V(x)\mathrm{d}H(x) = \int_0^1 \max\{V(x) : x \in [\overline{H}^{-1}(p), \underline{H}^{-1}(p)]\}\mathrm{d}p. \tag{2}$$

Conceptually, Theorem 2 follows easily from Theorem 1. Since the median matching experiment $\tau^\star$ implements all implementable distributions of medians, optimization just requires selecting an optimal median $\chi(G_\theta) \in \arg\max_{x\in[\theta,1/2+\theta]} V(x)$ for each posterior G_θ induced by $\tau^\star$, as illustrated in Figure 1. The value of the maximization problem is thus $2\int_0^{1/2} \max_{x\in[\theta,1/2+\theta]} V(x)\mathrm{d}\theta$, and a distribution H of medians is optimal iff $H^{-1}(2\theta) \in \arg\max_{x\in[\theta,1/2+\theta]} V(x)$ for all $\theta \in [0,1/2]$. That is, optimal solutions can be obtained by pointwise maximization without any ironing procedure.

In general, by Theorem 1, for each $H \in \mathcal{H}$ and $p \in [0,1]$, we have $\overline{H}^{-1}(p) \leq H^{-1}(p) \leq \underline{H}^{-1}(p)$. If we consider the relaxed problem of finding a measurable function $J : [0,1] \to \Theta$ to

$$\text{maximize} \int_0^1 V(J(p))\mathrm{d}p$$

[7] To establish results similar to our Theorem 1, Yang and Zentefis characterize the extreme points of the set $\{H \in \Delta(\Theta) : \underline{H} \leq H \leq \overline{H}\}$. As recently emphasized by Kleiner et al. (2021), characterizing a convex set by its extreme points can be useful for establishing a given property of the set. In contrast, we show that directly characterizing the set of implementable distributions of posterior quantiles is much easier than characterizing the extreme points of this set.

[8] Kolotilin et al. (2024) study the question of characterizing optimal distributions of such intermediate statistics—the analogous problem to that of Theorem 2 in the current paper.

$$\text{subject to } \overline{H}^{-1}(p) \leq J(p) \leq \underline{H}^{-1}(p), \quad \text{for all } p \in [0,1],$$

one solution is

$$J^\star(p) = \min \arg\max\{V(x) : x \in [\overline{H}^{-1}(p), \underline{H}^{-1}(p)]\}, \quad \text{for all } p \in [0,1].$$

This function $J^\star$ is monotone; moreover, the proof of Theorem 2 shows that there exists $H^\star \in \Delta(\Theta)$ such that $J^\star = H^{\star -1}$, so $H^\star$ solves the optimization problem (2).

The closest antecedent to Theorem 2 is Corollary 4 of Yang and Zentefis (2024), which solves the maximization problem (2) in the special cases where V is quasi-concave or quasi-convex. The solution follows immediately from Theorem 2. To see how, suppose that V is quasi-concave with a maximum at $x^\star \in [0,1]$. For each interval $[\theta, 1/2+\theta]$, it is optimal to select $x^\star$ if $x^\star \in [\theta, 1/2+\theta]$, θ if $x^\star < \theta$, and $1/2+\theta$ if $x^\star > 1/2+\theta$. This induces the distribution of posterior medians

$$H(x) = \begin{cases} \underline{H}(x), & x < x^\star, \\ \overline{H}(x), & x \geq x^\star. \end{cases}$$

Next, suppose that V is quasi-convex with $V(x^\star) = V(1/2 + x^\star)$ for some $x^\star \in [0, 1/2]$. Then, for each interval $[\theta, 1/2+\theta]$, it is optimal to select θ if $x^\star > \theta$ and $1/2+\theta$ if $x^\star < \theta$. This induces the distribution of posterior medians

$$H(x) = \begin{cases} \overline{H}(x), & x < x^\star, \\ 2x^\star, & x \in [x^\star, \frac{1}{2} + x^\star), \\ \underline{H}(x), & x \geq \frac{1}{2} + x^\star. \end{cases}$$

From the perspective of optimization, it is natural to ask whether each extreme point of $\mathcal{H}$ is exposed, meaning that it is the unique maximizer in $\mathcal{H}$ of $\int V(x)\mathrm{d}H(x)$ for some $V \in C(\Theta)$. It turns out that some extreme points are not exposed. To see this, note that in the uniform-median case the distribution $H^\star = (\delta_{1/4} + \delta_{1/2})/2$ is an extreme point of $\mathcal{H}$, as there are no distinct $H_1, H_2 \in \mathcal{H}$ such that $H^\star = (H_1 + H_2)/2$. By Theorem 2, if $H^\star$ maximizes $\int V(x)\mathrm{d}H(x)$ on $\mathcal{H}$ for some $V \in C(\Theta)$, then $V(1/4) \geq V(x)$ for all $x \in [\theta, 1/2+\theta]$ and all $\theta \in [0, 1/4]$, and similarly $V(1/2) \geq V(x)$ for all $x \in [\theta, 1/2+\theta]$ and all $\theta \in [0, 1/2]$. Thus, $V(1/4) = V(1/2) \geq V(x)$ for all $x \in [0,1]$. But then the distribution $\delta_{1/2} \in \mathcal{H}$ also maximizes $\int V(x)\mathrm{d}H(x)$, which shows that $H^\star$ is not an exposed point of $\mathcal{H}$.[9]

4. Unique properties of the quantile matching experiment

Theorem 1 shows that the q-quantile matching experiment simultaneously implements all implementable distributions of posterior q-quantiles. We now show that it is the

[9] The distribution $H^\star$ does uniquely maximize $\int V(x)\mathrm{d}H(x)$ for $V = 2 \cdot \mathbf{1}\{x = 1/4\} + \mathbf{1}\{x = 1/2\}$, which is upper semi-continuous but not continuous. An open question is whether each extreme point of $\mathcal{H}$, characterized in Theorem 1 of Yang and Zentefis (2024), is the unique maximizer of $\int V(x)\mathrm{d}H(x)$ for some upper-semicontinuous V. This is a weaker property than exposedness, as the usual theory of exposed points relies on continuity.

unique experiment to do so. For simplicity, in this section we assume that F has a positive density on Θ.

We actually establish the stronger result that the q-quantile matching experiment is the unique experiment that simultaneously implements all optimal distributions for strictly quasi-convex objective functions.

THEOREM 3. *The q-quantile matching experiment $\tau^\star$ is the unique experiment τ that, for each $p \in [0,1]$, implements the distribution $H_p \in \mathcal{H}$ given by*

$$H_p(x) = \begin{cases} \overline{H}(x), & x < \underline{x}_p, \\ p, & x \in [\underline{x}_p, \overline{x}_p), \\ \underline{H}(x), & x \geq \overline{x}_p, \end{cases}$$

where $\underline{x}_p = F^{-1}(qp)$ and $\overline{x}_p = F^{-1}(q + (1-q)p)$.

In other words, for any experiment $\tau \neq \tau^\star$, there is some $p \in [0,1]$ such that τ does not implement H_p. For example, in the uniform-median case, the negative assortative matching experiment does not implement $H_{1/2}$, as noted in the introduction.

An immediate corollary of Theorem 3 is that the q-quantile matching experiment is the unique experiment that minimizes the maximum regret of a designer who chooses an experiment τ before learning her objective V, but chooses a selection χ after learning V. Formally, for each experiment $\tau \in \Delta(\Delta(\Theta))$ and each objective $V \in C(\Theta)$ define the designer's *regret* as

$$r(\tau, V) = \max_{H \in \mathcal{H}} \int_\Theta V(x)\mathrm{d}H(x) - \sup_{H \in \mathcal{H}} \left\{ \int_\Theta V(x)\mathrm{d}H(x) : H \text{ is implemented by } \tau \right\}.$$

Note that $r(\tau, V) \geq 0$ for all τ and V. Say that a set of possible objective functions $\mathcal{V} \subset C(\Theta)$ is *rich* if, for all $x_0, x_1 \in \Theta$, there exists a strictly quasi-convex $V \in \mathcal{V}$ with $V(x_0) = V(x_1)$. We then have the following result.

COROLLARY 1. *If $\mathcal{V}$ is rich then the q-quantile matching experiment $\tau^\star$ is the unique experiment τ such that $r(\tau, V) = 0$ for all $V \in \mathcal{V}$.*

APPENDIX: PROOFS

PROOF OF THEOREM 1. Consider any experiment $\tau \in \Delta(\Delta(\Theta))$ and any measurable selection $\chi(G)$ from $X(G)$. Let H be the distribution of $\chi(G)$ induced by τ. Then, for each $x \in \Theta$, we have

$$\begin{aligned} F(x) = \int G(x)\mathrm{d}\tau(G) &= \int \mathbf{1}\{G(x) \geq q\}G(x)\mathrm{d}\tau(G) + \int \mathbf{1}\{G(x) < q\}G(x)\mathrm{d}\tau(G) \\ &\geq \int \mathbf{1}\{G(x) \geq q\}q\mathrm{d}\tau(G) \geq \int \mathbf{1}\{\chi(G) \leq x\}q\mathrm{d}\tau(G) = qH(x), \end{aligned}$$

showing that $H \leq \overline{H}$. A symmetric argument shows that $H \geq \underline{H}$.

For the converse, we first note that the median matching experiment $\tau^\star$ is well-defined because $\int G \mathrm{d}\tau^\star(G) = F$: indeed, for all $\theta \in \Theta$, we have

$$\int G(\theta)\mathrm{d}\tau^\star(G) = \int_0^q \left(q\delta_{F^{-1}(\omega)} + (1-q)\delta_{F^{-1}(q+\frac{1-q}{q}\omega)} \right)(\theta)\,\frac{\mathrm{d}\omega}{q}$$
$$= \begin{cases} \int_0^{F(\theta)} q\frac{\mathrm{d}\omega}{q}, & F(\theta) < q, \\ \int_0^q q\frac{\mathrm{d}\omega}{q} + \int_0^{\frac{q}{1-q}(F(\theta)-q)} (1-q)\frac{\mathrm{d}\omega}{q}, & F(\theta) \geq q, \end{cases} = F(\theta),$$

where the second equality holds because $F^{-1}(\omega) \leq \theta$ iff $\omega \leq F(\theta)$, and $F^{-1}(q+\frac{1-q}{q}\omega) \leq \theta$ iff $\omega \leq \frac{q}{1-q}(F(\theta)-q)$. Note also that, for each $\omega \in [0,q]$, the set of q-quantiles of G_ω is $X(G_\omega) = [F^{-1}(\omega), F^{-1}(q+\frac{1-q}{q}\omega)]$.

Now fix a distribution $H \in \Delta(\Theta)$ satisfying $\underline{H} \leq H \leq \overline{H}$. Note that, for each $\omega \in [0,q]$, since $H \leq \overline{H}$, we have $H^{-1}(\frac{\omega}{q}) \geq \overline{H}^{-1}(\frac{\omega}{q}) = F^{-1}(\omega)$; and, since $H \geq \underline{H}$, we have $H^{-1}(\frac{\omega}{q}) \leq \underline{H}^{-1}(\frac{\omega}{q}) = F^{-1}(q+\frac{1-q}{q}\omega)$. Thus, $H^{-1}(\frac{\omega}{q}) \in X(G_\omega)$. We can therefore define a selection $\chi(G)$ from $X(G)$ by letting $\chi(G) = H^{-1}(\frac{\omega}{q})$ in the $\tau^\star$-almost sure event that $G = G_\omega$ for some $\omega \in [0,q]$, and (for concreteness) letting $\chi(G) = \min X(G)$ otherwise. Finally, the distribution of $\chi(G)$ induced by $\tau^\star$ is H, because, for all $x \in \Theta$, we have

$$\int \mathbf{1}\{\chi(G) \leq x\}\mathrm{d}\tau^\star(G) = \int \mathbf{1}\{H^{-1}(\tfrac{\omega}{q}) \leq x\}\frac{\mathrm{d}\omega}{q} = \int_0^{qH(x)} \frac{\mathrm{d}\omega}{q} = H(x).$$

For unique implementation, assume that F has a positive density on $\Theta = [\underline{\theta}, \overline{\theta}]$. Fix any $H \in \mathcal{H}$. Consider a sequence of partitions of Θ given by $\theta_{i,n} = \underline{\theta} + (\overline{\theta} - \underline{\theta})\frac{i}{2^n}$, with $i \in \{0,1,\ldots,2^n\}$. Define a sequence $H_n \in \Delta(\Theta)$ by

$$H_n(x) = H(\theta_{i-1,n})\frac{F(\theta_{i,n}) - F(x)}{F(\theta_{i,n}) - F(\theta_{i-1,n})} + H(\theta_{i,n})\frac{F(x) - F(\theta_{i-1,n})}{F(\theta_{i,n}) - F(\theta_{i-1,n})},$$

for all $i \in \{1,\ldots,2^n\}$ and all $x \in [\theta_{i-1,n}, \theta_{i,n}]$. Note that H_n is well-defined, because F is strictly increasing on Θ. Since $H \in \mathcal{H}$, we have $H_n \in \mathcal{H}$. Moreover, H_n has a simple density function h_n with respect to F, given by

$$h_n(x) = \frac{H(\theta_{i,n}) - H(\theta_{i-1,n})}{F(\theta_{i,n}) - F(\theta_{i-1,n})},$$

for all $i \in \{1,\ldots,2^n\}$ and all $x \in (\theta_{i-1,n}, \theta_{i,n})$.

Next, for each $e \in (0,1]$ and each n, there exists an experiment $\tau^\star_{e,n} \in \Delta(\Delta(\Theta))$ satisfying the following two properties. First, for $\tau^\star_{e,n}$-almost all G, there exists $x \in \Theta$ such that $G = G^x$ where

$$G^x = \frac{(1-e)h_n(x)(q\delta_{F^{-1}(qH_n(x))} + (1-q)\delta_{F^{-1}(q+(1-q)H_n(x))}) + e\delta_x}{(1-e)h_n(x) + e}.$$

This implies that $X(G^x)$ is the singleton $\{x\}$, because $e > 0$ and $F^{-1}(qH_n(x)) \leq x \leq F^{-1}(q+(1-q)H_n(x))$ (which holds because $\underline{H}(x) \leq H_n(x) \leq \overline{H}(x)$). Second, the distribution of unique quantiles $\chi(G^x) = x$ induced by $\tau^\star_{e,n}$ is $(1-e)H_n + eF$.

Formally, $\tau^\star_{e,n}$ is defined by

$$\tau^\star_{e,n}(M) = \int_0^1 \mathbf{1}\{G^x \in M\}\,((1-e)h_n(x)+e)\,\mathrm{d}F(x), \quad \text{for all } M \subset \Delta(\Theta).$$

Note that $\tau^\star_{e,n}$ is a well-defined experiment because $\int G \mathrm{d}\tau^\star_{e,n}(G) = F$. Indeed, for all $\theta \in \Theta$, we have

$$\int G(\theta)\mathrm{d}\tau^\star_{e,n}(G) = (1-e)\int_0^1 h_n(x)(q\delta_{F^{-1}(qH_n(x))} + (1-q)\delta_{F^{-1}(q+(1-q)H_n(x))})\,(\theta)\,\mathrm{d}F(x)$$

$$+\, e\int_0^1 \delta_x(\theta)\,\mathrm{d}F(x) = (1-e)\begin{cases} \int_0^{H_n^{-1}(\frac{F(\theta)}{q})} q\mathrm{d}H_n(x), & F(\theta) < q, \\ \int_0^{H_n^{-1}(1)} q\mathrm{d}H_n(x) + \int_0^{H_n^{-1}(\frac{F(\theta)-q}{1-q})}(1-q)\mathrm{d}H_n(x), & F(\theta) \geq q, \end{cases}$$

$$+\, e\int_0^\theta \mathrm{d}F(x) = F(\theta),$$

where the second equality holds because $F^{-1}(qH_n(x)) \leq \theta$ iff $x \leq H_n^{-1}(\frac{F(\theta)}{q})$, and $F^{-1}(q+(1-q)H_n(x)) \leq \theta$ iff $x \leq H_n^{-1}(\frac{F(\theta)-q}{1-q})$; and the third equality holds because $H_n(H_n^{-1}(p)) = p$ for all $p \in [0,1]$, by continuity of H_n (which holds because H_n has a density with respect to F and F has a density with respect to the Lebesgue measure). Finally, the distribution of unique quantiles $\chi(G)$ induced by $\tau^\star_{e,n}$ is $(1-e)H_n + eF$, because, for all $y \in \Theta$, we have

$$\int \mathbf{1}\{\chi(G) \leq y\}\mathrm{d}\tau^\star_{e,n}(G) = \int_0^y ((1-e)h_n(x)+e)\mathrm{d}F(x) = (1-e)H_n(y) + eF(y).$$

Now fix $V \in C(\Theta)$. By continuity of V and compactness of Θ, for all $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that (i) $|V(x) - V(y)| \leq \varepsilon$ for all $x, y \in [\theta_{i-1,n}, \theta_{i,n}]$, all $i \in \{1, \ldots, 2^n\}$, and all $n \geq N$, and (ii) $e|V(x) - V(y)| \leq \varepsilon$ for all $x, y \in \Theta$ and all $e \in (0, \frac{1}{N}]$. Then, for all $n \geq N$ and $e \in (0, \frac{1}{N}]$, we have

$$\left|\int V(x)\mathrm{d}H(x) - \int V(x)\mathrm{d}((1-e)H_n + eF)(x)\right|$$
$$\leq (1-e)\left|\int V(x)\mathrm{d}(H-H_n)(x)\right| + e\left|\int V(x)\mathrm{d}(H-F)(x)\right| \leq \varepsilon + \varepsilon.$$

Since this holds for any $V \in C(\Theta)$, it follows that $(1-\frac{1}{n})H_n + \frac{1}{n}F$ converges weakly to H. In turn, since we have seen that the experiment $\tau^\star_{1/n,n}$ uniquely implements $(1-\frac{1}{n})H_n + \frac{1}{n}F$, we conclude that $\mathcal{H}^\star$ is dense in $\mathcal{H}$. Finally, $\mathcal{H}$ is compact, as $\Delta(\Theta)$ is compact by Theorem 15.11 in Aliprantis and Border (2006), and $\mathcal{H}$ is the intersection over $x \in \Theta$ of the closed subsets $\mathcal{H}_x := \{H \in \Delta(\Theta) : \underline{H}(x) \leq H(x) \leq \overline{H}(x)\}$ of $\Delta(\Theta)$. Thus, the closure of $\mathcal{H}^\star$ is $\mathcal{H}$, and hence (1) holds for any $V \in C(\Theta)$. □

PROOF OF THEOREM 2. For each $H \in \Delta(\Theta)$, we have

$$\int_\Theta V(x)\mathrm{d}H(x) = \int_0^1 V(H^{-1}(p))\mathrm{d}p.$$

Recall that

$$J^\star(p) = \min \arg\max\{V(x) : x \in [\overline{H}^{-1}(p), \underline{H}^{-1}(p)]\}, \quad \text{for all } p \in [0,1].$$

Since $J^\star$ is defined as the minimum selection from the $\arg\max$, it follows that (i) $J^\star$ is non-decreasing (and hence measurable), because $\overline{H}^{-1}$ and $\underline{H}^{-1}$ are non-decreasing, (ii) $J^\star$ is left-continuous, because $\overline{H}^{-1}$ and $\underline{H}^{-1}$ are left-continuous and $V \in C(\Theta)$, (iii) $J^\star(1) \le \overline{\theta}$, because $\underline{H}^{-1}(1) \le \overline{\theta}$, and (iv) $J^\star(0) = \underline{\theta}$, because $\overline{H}^{-1}(0) = \underline{H}^{-1}(0) = \underline{\theta}$. This implies that $J^\star = H^{\star -1}$, where $H^\star \in \Delta(\Theta)$ is given by

$$H^\star(x) = \sup\{p \in [0,1] : J^\star(p) \le x\}, \quad \text{for all } x \in \Theta.$$

Moreover, since $\overline{H}^{-1} \le J^\star \le \underline{H}^{-1}$, it follows that $H^\star \in \mathcal{H}$, so $H^\star$ solves the original problem, and its value coincides with the value of the relaxed problem, yielding (2). Consequently, $H \in \mathcal{H}$ maximizes $\int V(x)\mathrm{d}H(x)$ on $\mathcal{H}$ iff $H^{-1}(p)$ maximizes V on $[\overline{H}^{-1}(p), \underline{H}^{-1}(p)]$ for almost all $p \in [0,1]$. Moreover, by continuity of V and left-continuity of $\overline{H}^{-1}$ and $\underline{H}^{-1}$, $H^{-1}(p)$ maximizes V on $[\overline{H}^{-1}(p), \underline{H}^{-1}(p)]$ for almost all $p \in [0,1]$ iff it does so for all $p \in [0,1]$.

Next, if V has a unique maximum on $[\overline{H}^{-1}(p), \underline{H}^{-1}(p)]$ for almost all $p \in [0,1]$, then $J^\star$ is the unique solution of the relaxed problem that satisfies properties (i)–(iv), and hence $H^\star$ is the unique solution of the original problem. Conversely, if there exists a non-negligible set $P \subset [0,1]$ such that V has multiple maxima on $[\overline{H}^{-1}(p), \underline{H}^{-1}(p)]$ for each $p \in P$, then there are multiple solutions of the relaxed problem that satisfy properties (i)–(iv). For example, $\hat{J}$ defined as the maximum selection from the $\arg\max$ also solves the relaxed problem, and so does $\hat{J}^\star$ defined by $\hat{J}^\star(p) = \hat{J}(p^-)$ for all $p \in (0,1]$ and $\hat{J}^\star(0) = \underline{\theta}$. But, by construction, $\hat{J}^\star$ satisfies properties (i)–(iv) and is not equal to $J^\star$. Then $\hat{J}^\star = \hat{H}^{\star -1}$ where $\hat{H}^\star \in \Delta(\Theta)$ is given by

$$\hat{H}^\star(x) = \sup\{p \in [0,1] : \hat{J}^\star(p) \le x\}, \quad \text{for all } x \in \Theta.$$

Thus, $\hat{H}^\star \neq H^\star$ also solves the original problem. □

Proof of Theorem 3. Suppose that an experiment $\tau \in \Delta(\Delta(\Theta))$ implements all H_p. Fix any $p \in [0,1]$. Since τ implements H_p, there exists a measurable selection $\chi_p(G)$ from $X(G)$ such that the distribution of $\chi_p(G)$ induced by τ is H_p. Since F has a density on Θ, we have

$$\begin{aligned} qp = F(\underline{x}_p) = \int G(\underline{x}_p)\mathrm{d}\tau(G) = \int \mathbf{1}\{G(\underline{x}_p) \ge q\} G(\underline{x}_p)\mathrm{d}\tau(G) \\ + \int \mathbf{1}\{G(\underline{x}_p) < q\} G(\underline{x}_p)\mathrm{d}\tau(G) \ge \int \mathbf{1}\{G(\underline{x}_p) \ge q\} q\mathrm{d}\tau(G) \\ \ge \int \mathbf{1}\{\chi_p(G) \le \underline{x}_p\} q\mathrm{d}\tau(G) = qH_p(\underline{x}_p) = qp, \end{aligned}$$

so all inequalities hold with equality. Thus, $\tau(G(\underline{x}_p) = 0) = 1 - p$, $\tau(G(\underline{x}_p) = q) = p$, and $\tau(\chi_p(G) \le \underline{x}_p) = p$. A symmetric argument yields $\tau(G(\overline{x}_p) = q) = 1 - p$, $\tau(G(\overline{x}_p) = 1) = p$, and $\tau(\chi_p(G) > \overline{x}_p) = 1 - p$. Next, since $G(\underline{x}_p) = 0$ and $G(\overline{x}_p) = 1$ imply that $\underline{x}_p < \chi_p(G) \leq \overline{x}_p$, it follows that $\tau(G(\underline{x}_p) = 0,\, G(\overline{x}_p) = 1) = 0$, because

$$\tau(\underline{x}_p < \chi_p(G) \leq \overline{x}_p) = 1 - \tau(\chi_p(G) \leq \underline{x}_p) - \tau(\chi_p(G) > \overline{x}_p) = 1 - p - (1-p) = 0.$$

So, $\tau(G(\underline{x}_p) = 0,\, G(\overline{x}_p) = q) = \tau(G(\underline{x}_p) = 0) - \tau(G(\underline{x}_p) = 0,\, G(\overline{x}_p) = 1) = 1 - p$ and $\tau(G(\underline{x}_p) = q,\, G(\overline{x}_p) = 1) = \tau(G(\overline{x}_p) = 1) - \tau(G(\underline{x}_p) = 0,\, G(\overline{x}_p) = 1) = p$. In sum,

$$\begin{aligned} &\tau(G(\underline{x}_p) = 0,\, G(\overline{x}_p) = q) = 1 - p, \\ &\tau(G(\underline{x}_p) = q,\, G(\overline{x}_p) = 1) = p, \end{aligned} \qquad \text{for all } p \in [0,1]. \tag{3}$$

We now show that (3) yields $\tau = \tau^\star$. Let $X_0 = [\underline{x}_0, \overline{x}_0] = [\underline{\theta}, F^{-1}(q)]$ and $X_1 = [\underline{x}_1, \overline{x}_1] = [F^{-1}(q), \overline{\theta}]$. For each experiment $\tilde{\tau} \in \Delta(\Delta(\Theta))$, define a joint distribution function $I_{\tilde{\tau}} : X_0 \times X_1 \to [0,1]$ by

$$I_{\tilde{\tau}}(x_0, x_1) = \tilde{\tau}(G = q\delta_{\theta_0} + (1-q)\delta_{\theta_1},\, \theta_0 \in [\underline{x}_0, x_0],\, \theta_1 \in [\underline{x}_1, x_1]).$$

To prove that $\tau = \tau^\star$, it suffices to show that $I_\tau(x_0, x_1) = I_{\tau^\star}(x_0, x_1)$ for all $(x_0, x_1) \in X_0 \times X_1$, with $I_\tau(\overline{x}_0, \overline{x}_1) = I_{\tau^\star}(\overline{x}_0, \overline{x}_1) = 1$. Fix any $(x_0, x_1) \in X_0 \times X_1$, and let $\hat{p} = \min\{\frac{F(x_0)}{q}, \frac{F(x_1) - q}{1-q}\}$. First, by definition of $\tau^\star$, we have

$$I_{\tau^\star}(x_0, x_1) = \int_0^q \mathbf{1}\{F^{-1}(\omega) \leq x_0,\, F^{-1}(q + \tfrac{1-q}{q}\omega) \leq x_1\} \frac{\mathrm{d}\omega}{q} = \hat{p},$$

with $I_{\tau^\star}(\overline{x}_0, \overline{x}_1) = 1$, because $F(\overline{x}_0) = F(F^{-1}(q)) = q$ and $F(\overline{x}_1) = F(\overline{\theta}) = 1$. Second, by (3) and definition of I_τ, we have

$$\hat{p} = \tau(G(\underline{x}_{\hat{p}}) = q,\, G(\overline{x}_{\hat{p}}) = 1) \leq I_\tau(x_0, x_1) \leq \tau((G(\underline{x}_{\hat{p}}), G(\overline{x}_{\hat{p}})) \neq (0, q)) = 1 - (1 - \hat{p}),$$

showing that $I_\tau(x_0, x_1) = \hat{p} = I_{\tau^\star}(x_0, x_1)$. □

PROOF OF COROLLARY 1. Consider any experiment $\tau \neq \tau^\star$. By Theorem 3, there exists $p \in [0,1]$ such that τ does not implement H_p. Since $\mathcal{V}$ is rich, there exists a continuous and strictly quasi-convex $V \in \mathcal{V}$ with $V(\underline{x}_p) = V(\overline{x}_p)$. Then $\underline{x}_{\tilde{p}}$ uniquely maximizes V on $[\underline{x}_{\tilde{p}}, \overline{x}_{\tilde{p}}]$ for all $\tilde{p} \in [0, p)$, and $\overline{x}_{\tilde{p}}$ uniquely maximizes V on $[\underline{x}_{\tilde{p}}, \overline{x}_{\tilde{p}}]$ for all $\tilde{p} \in (p, 1]$. By Theorem 2, H_p uniquely maximizes $\int V(x)\mathrm{d}H(x)$ on $\mathcal{H}$.

Suppose for contradiction that $r(\tau, V) = 0$. Then there exists a sequence $H^n \in \mathcal{H}$ implemented by τ such that $\int V(x)\mathrm{d}H^n(x) \to \int V(x)\mathrm{d}H_p(x)$. Since $\mathcal{H}$ is weak* compact, passing to a subsequence if necessary, we can assume that $H^n \to \hat{H} \in \mathcal{H}$. Note that $\hat{H} = H_p$, because H_p uniquely maximizes $\int V(x)\mathrm{d}H(x)$ on $\mathcal{H}$. In sum, there exists a sequence of measurable selections $\chi^n(G)$ from $X(G)$ such that $H^n(x) = \tau(\chi^n(G) \leq x)$ and $H^n(x) \to H_p(x)$ for all $x \in \Theta$.

Next, we show that τ cannot simultaneously satisfy the following three conditions

$$\tau(G(\underline{x}_{\tilde{p}}) = 0) = 1 - \tilde{p} \quad \text{and} \quad \tau(G(\underline{x}_{\tilde{p}}) = q) = \tilde{p}, \quad \text{for all } \tilde{p} \in [0, p], \tag{4}$$

$$\tau(G(\overline{x}_{\tilde{p}}) = q) = 1 - \tilde{p} \quad \text{and} \quad \tau(G(\overline{x}_{\tilde{p}}) = 1) = \tilde{p}, \quad \text{for all } \tilde{p} \in [p, 1], \tag{5}$$

$$\tau(G(\underline{x}_p)=0,\,G(\overline{x}_p)=q)=1-p \quad \text{and} \quad \tau(G(\underline{x}_p)=q,\,G(\overline{x}_p)=1)=p, \tag{6}$$

as otherwise τ would implement H_p. Indeed, if τ satisfies (4)–(6), we can define a selection $\chi(G)$ from $X(G)$ by letting $\chi(G)=\theta_0$ if $\theta_0<\underline{x}_p$ and $\chi(G)=\theta_1$ if $\theta_0>\underline{x}_p$ in the τ-almost sure event that $G=q\delta_{\theta_0}+(1-q)\delta_{\theta_1}$ for some $\theta_0\in[\underline{x}_0,\overline{x}_0]$ and $\theta_1\in[\underline{x}_1,\overline{x}_1]$. Then the distribution of $\chi(G)$ induced by τ is H_p, because, by (4), for all $\tilde{p}\le p$, we have

$$\tau(\chi(G)\le\underline{x}_{\tilde{p}})=\tau(G(\underline{x}_{\tilde{p}})=q)=\tilde{p}=\tfrac{F(\underline{x}_{\tilde{p}})}{q}=H_p(\underline{x}_{\tilde{p}}),$$

and, by (5) and (6), for all $\tilde{p}\ge p$, we have

$$\tau(\chi(G)\le\overline{x}_{\tilde{p}})=\tau(G(\overline{x}_{\tilde{p}})=1)+\tau(G(\underline{x}_p)=q,\,G(\overline{x}_{\tilde{p}})=q)=\tilde{p}=\tfrac{F(\overline{x}_{\tilde{p}})-q}{1-q}=H_p(\overline{x}_{\tilde{p}}).$$

Finally, we show that if at least one of conditions (4)–(6) fails, then $H^n\nrightarrow H_p$. First, if (4) fails at some $\tilde{p}\in[0,p]$, then there exists $\varepsilon>0$ such that

$$\begin{aligned}F(\underline{x}_{\tilde{p}})&=\int G(\underline{x}_{\tilde{p}})\mathrm{d}\tau(G)=\int \mathbf{1}\{G(\underline{x}_{\tilde{p}})\ge q\}G(\underline{x}_{\tilde{p}})\mathrm{d}\tau(G)+\int \mathbf{1}\{G(\underline{x}_{\tilde{p}})<q\}G(\underline{x}_{\tilde{p}})\mathrm{d}\tau(G)\\&\ge\varepsilon+\int \mathbf{1}\{G(\underline{x}_{\tilde{p}})\ge q\}q\mathrm{d}\tau(G)\ge\varepsilon+\int \mathbf{1}\{\chi^n(G)\le\underline{x}_{\tilde{p}}\}q\mathrm{d}\tau(G)=\varepsilon+qH^n(\underline{x}_{\tilde{p}}),\end{aligned}$$

so $H^n(\underline{x}_{\tilde{p}})\nrightarrow H_p(\underline{x}_{\tilde{p}})$. Similarly, if (5) fails at some $\tilde{p}\in[p,1]$, then $H^n(\overline{x}_{\tilde{p}})\nrightarrow H_p(\overline{x}_{\tilde{p}})$. Finally, if (4) and (5) hold, but (6) fails, then there exists $\varepsilon>0$ such that

$$\begin{aligned}H^n(\overline{x}_p)-H^n(\underline{x}_p)&=\tau(\underline{x}_p<\chi^n(G)\le\overline{x}_p)\ge\tau(G(\underline{x}_p)=0,\,G(\overline{x}_p)=1)\\&\ge\varepsilon>0=H^p(\overline{x}_p)-H^p(\underline{x}_p),\end{aligned}$$

so $H^n(\underline{x}_p)\nrightarrow H_p(\underline{x}_p)$ or $H^n(\overline{x}_p)\nrightarrow H_p(\overline{x}_p)$. □

References

Aliprantis, Charalambos D. and Kim Border (2006), *Infinite Dimensional Analysis: A Hitchhiker's Guide*. Springer. [10]

Benoît, Jean-Pierre and Juan Dubra (2011), "Apparent overconfidence." *Econometrica*, 79, 1591–1625. [1, 3, 5, 6]

Blackwell, David (1953), "Equivalent comparisons of experiments." *Annals of Mathematical Statistics*, 24, 265–272. [6]

Friedman, John N and Richard T Holden (2008), "Optimal gerrymandering: Sometimes pack, but never crack." *American Economic Review*, 98(1), 113–44. [1, 5]

Kamenica, Emir and Matthew Gentzkow (2011), "Bayesian persuasion." *American Economic Review*, 101, 2590–2615. [2, 4, 5]

Kleiner, Andreas, Benny Moldovanu, and Philipp Strack (2021), "Extreme points and majorization: Economic applications." *Econometrica*, 89(4), 1557–1593. [6]

Kolotilin, Anton (2018), "Optimal information disclosure: A linear programming approach." *Theoretical Economics*, 13, 607–636. [6]

Kolotilin, Anton, Roberto Corrao, and Alexander Wolitzky (2024), "Persuasion and matching: Optimal productive transport." *Journal of Political Economy*, forthcoming. [1, 6]

Kolotilin, Anton and Alexander Wolitzky (2020a), "Assortative information disclosure." *SSRN Working Paper*, https://ssrn.com/abstract=3616781. [1, 2, 6]

Kolotilin, Anton and Alexander Wolitzky (2020b), "The economics of partisan gerrymandering." *SSRN Working Paper*, https://ssrn.com/abstract=3698690. [1, 6]

Kremer, Michael and Eric Maskin (1996), "Wage inequality and segregation by skill." *NBER Working Paper*, https://www.nber.org/papers/w5718. [2]

Legros, Patrick and Andrew F Newman (2002), "Monotone matching in perfect and imperfect worlds." *Review of Economic Studies*, 69(4), 925–942. [2]

Owen, Guillermo and Bernard Grofman (1988), "Optimal partisan gerrymandering." *Political Geography Quarterly*, 7(1), 5–22. [2]

Strassen, Volker (1965), "The existence of probability measures with given marginals." *Annals of Mathematical Statistics*, 36, 423–439. [6]

Yang, Kai Hao and Alexander Zentefis (2024), "The economics of monotone function intervals." *American Economic Review*, forthcoming. [1, 2, 3, 6, 7]