Distributions of Posterior Quantiles and Economic Applications*

Kai Hao Yang† Alexander K. Zentefis‡

July 7, 2022

Abstract

We characterize the distributions of posterior quantiles under a given prior. Unlike the distributions of posterior means, which are known to be mean-preserving contractions of the prior, the distributions of posterior quantiles coincide with a first-order stochastic dominance interval bounded by an upper and a lower truncation of the prior. We apply this characterization to several environments, ranging across political economy, Bayesian persuasion, industrial organization, econometrics, finance, and accounting.

JEL classification: C31, D72, D82, D83, G28, L12, M41

Keywords: Posterior quantiles, stochastic dominance, gerrymandering, Bayesian persuasion, two-sided markets, quantile regression

*This paper subsumes a previously circulated paper with the title "Gerrymandering and the Limits of Representative Democracy." We are grateful to Yuehao Bai, Ian Ball, Dirk Bergemann, Ben Brooks, Peter Caradonna, Roberto Corrao, Kuan-ming Chen, Xiaohong Chen, Yu-ting Chiang, Joyee Deb, Piotr Dworczak, Tan Gan, Paul Goldsmith-Pinkham, Yannai Gonczarowski, Marina Halac, Ángel Hernando-Veciana, Ryota Iijima, Emir Kamenica, Alexey Kushnir, Elliot Lipnowski, Leslie Marx, Stephen Morris, Barry Nalebuff, Aniko Öry, Wolfgang Pesendorfer, Benjamin Polak, Doron Ravid, Anne-Katrin Roesler, Kareen Rozen, Hamid Sabourian, Chris Shannon, Quitzé Valenzuela-Stookey, Vasiliki Skreta, Nicholas Stephanopoulos, Philipp Strack, Roland Strausz, Tomasz Strzalecki, Alexander Wolitzky, Wenji Xu, Nathan Yoder, Jidong Zhou; and the participants of various seminars and conferences for their helpful comments. All errors are our own.

†Yale School of Management, Email: kaihao.yang@yale.edu

‡Yale School of Management, Email: alexander.zentefis@yale.edu

1 Introduction

A political body is redrawing the boundaries of electoral districts for partisan gain. Election results are dictated by the median voter theorem. To what extent can this gerrymandering affect the composition of the legislature?

A ride-sharing app is a platform between riders and drivers, and it can segment both sides of the market. An inelastic supply must be held fixed at 75% of the number of riders in each segment, so that wait times remain approximately the same across riders. This means that, in each segment, the top 75th percentile of riders' willingness-to-pay determines the price. Which segmentation maximizes the platform's revenue, and which one maximizes consumer surplus?

An econometrician observes data on income and education from two different samples of the population. What can she infer about the relation between the top 1% of earners and their years of schooling?

Despite these scenarios roaming varied economic fields; reaching the areas of political economy, industrial organization, and econometrics; all are connected by their shared concern over the distribution of different posterior quantiles. This paper characterizes the distributions of posterior quantiles in a general setting, and it answers each of the scenario's question and more.

In our environment, a one-dimensional variable $\omega \in \mathbb{R}$ follows a prior distribution F_0. Given any signal for ω (i.e., a joint distribution of ω and signal realizations with the marginal of ω being F_0), each signal realization induces a *posterior* distribution of ω via Bayes' rule. Therefore, any signal induces a distribution of posterior beliefs. If one computes the mean of each posterior, Strassen's theorem (Strassen 1965) implies that the distribution of posterior means must be a mean-preserving contraction of F_0. At the same time, any mean-preserving contraction of F_0 is a distribution of posterior means induced by some signal.

Instead of posterior means, one can derive many other statistics of a posterior. Suppose that, instead of the mean, one computes a τ-quantile. A natural question then follows: For any $\tau \in (0, 1)$, what is the distribution of posterior τ-quantiles?

Theorem 1 provides that characterization. Two distributions are important in this regard:

$$\underline{F}_0^\tau(\omega) := \min\left\{\frac{1}{\tau}F_0(\omega), 1\right\}, \quad \overline{F}_0^\tau(\omega) := \max\left\{\frac{F_0(\omega) - \tau}{1 - \tau}, 0\right\}.$$

The distribution $\underline{F}_0^\tau$ can be interpreted as the conditional distribution of F_0 in the event that ω is smaller than a τ-quantile of F_0. Similarly, $\overline{F}_0^\tau$ can be interpreted as the conditional distribution of F_0 in the event that ω is larger than the same τ-quantile. Theorem 1 states

that any distribution of posterior τ-quantiles induced by a signal must be bounded by $\underline{F}_0^\tau$ and $\overline{F}_0^\tau$ in the sense of first-order stochastic dominance. In the meantime, any distribution bounded by $\underline{F}_0^\tau$ and $\overline{F}_0^\tau$ in the sense of first-order stochastic dominance must be a distribution of posterior τ-quantiles under some signal.

With the characterization of the distributions of posterior quantiles in hand, we then apply it to several economic settings. The first is to gerrymandering, or the manipulation of electoral district boundaries. In this setting, citizens identify with an ideal position on political issues along a spectrum. The variety of positions is represented as a distribution, which we can call a prior. An electoral map segments citizens into districts, which splits the prior distribution into different parts. The distribution of ideal positions within each district can be interpreted as a posterior.

If each district elects a representative holding the district's median position (Downs 1957; Black 1958), the composition of the legislative body (i.e., the distribution of ideal positions of elected representatives) can then be represented by a distribution of posterior medians. In this regard, Theorem 1 fully describes the scope of legislatures that unrestrained gerrymandering can achieve. According to Theorem 1, gerrymandering can induce *any* legislature within the bounds of two extremes: an "all-left" body and an "all-right" body. In the former, every representative occupying the legislature has an ideal position that is left of the median voter's ideal, whereas in the latter, every representative is to the right. These two bounds imply that unrestricted gerrymandering, taken to the extreme, can induce an entirely one-sided congress. Theorem 1 also sheds light on how polarized a legislature can become. According to the theorem, unrestrained gerrymandering can lead to a chamber without a single "moderate" member whose ideal position is in the interquartile range of the prior.

With knowledge about possible compositions of a legislative body, we then study the effects of gerrymandering on enacted legislation. For any legislative voting procedure that enacts a policy whenever it is the ideal position of a majority of representatives (e.g., one that enacts a median of a legislature's views), an immediate consequence of Theorem 1 is that *any* position that is part of the interquartile range of the prior can be enacted by a legislature under some map. As a consequence, more extreme legislation can be enacted if the population becomes more polarized in their views.

In addition to gerrymandering, we apply our characterization to Bayesian persuasion. Kamenica and Gentzkow (2011) provide a framework for studying a sender's communication to a receiver under the commitment assumption. A practical challenge, however, is that the concavafication approach used in this literature loses tractability as the number of states increases. An exception is when the state is one-dimensional and only posterior means are payoff-relevant to the sender. Theorem 1 complements this literature, as it brings tractability

to settings where only posterior *quantiles* are payoff-relevant to the sender. We demonstrate this by revisiting the two examples of Kamenica and Gentzkow (2011). Even in settings without the commitment assumption, Theorem 1 still facilitates the characterization of the sender's equilibrium payoff. We demonstrate this by revisiting the cheap talk model with transparent motives from Lipnowski and Ravid (2020).

Bayesian persuasion has notably been applied to industrial organization settings involving market segmentation, in which a market is split into several segments to further price discrimination (e.g., Bergemann, Brooks, and Morris 2015; Ichihashi 2020). In our next application, we use the characterization to derive the outcomes induced by different segmentations of a two-sided market.

In our application, we consider a two-sided market (e.g., ride-sharing). The demand side is populated by riders with unit demand; whereas the supply side is populated by drivers with unit supply. Total supply is inelastic, which is plausible during peak hours at a major airport or at the conclusion of a large sporting event. A third-party platform (the ride-sharing app) can segment the market to affect prices, but it is obligated to keep the ratio of supply to demand of each segment the same (i.e., it faces a *market thickness* constraint), so that all riders wait approximately the same time before matching with a driver. Theorem 1 provides a characterization of all outcomes that two-sided market segmentation can induce. Perhaps surprisingly, if the platform's revenue depends solely on total sales, the platform's optimal segmentation fully extracts all consumer surplus, rendering any thickness constraint irrelevant.

Because the distribution of posterior quantiles is simply a conditional distribution, another natural discipline ready for applications is econometrics. We apply the characterization to quantile regression, which models the quantiles of the conditional distribution of a response variable Y as a function of covariates X.

To facilitate their analysis and maintain tractability, econometricians often impose parametric assumptions regarding the quantile function, such as presuming linearity. Theorem 1 provides a model mis-specification test that relies only on the marginal of Y. The reliance on information from just the marginal allows one to bypass estimation of the joint distribution of (Y, X), which may be computationally demanding.

When the number of covariates equals 1 and the marginals of both X and Y are known, an econometrician can go beyond evaluating model mis-specification to partially identifying the quantile function. Taking a concrete example, one might have Y representing income and X standing for education, and the two sets of data come from non-overlapping samples of the population. If the quantile function is known to be increasing (such as income rising with years of schooling), we show how Theorem 1 provides a non-parametric partial identification

of the quantile function.

Our final applications are to topics in finance and accounting. For the financial application, we take a setting in which a bank regulator considers requiring systemically important financial institutions to issue equity capital in accordance with their contribution to the Value-at-Risk of the financial system. Our characterization describes the full range of equity that financial institutions would have to raise if under duress. For accounting, we consider an auditor who worries about a manager misclassifying expenses to boost earnings. A certain dollar amount of misclassification constitutes material fraud. We provide a necessary condition for an auditor to engage in a closer inspection of the reported expenses.

Related Literature This paper relates to several areas. Belief-based characterizations of signals date back to the seminal contributions of Blackwell (1953) and Harsanyi (1967-68). The characterization of distributions of posterior means can be derived from Strassen (1965), and their economic applications are made clear in Rothschild and Stiglitz (1970). This paper can be regarded as a complement, as it characterizes the distributions of posterior quantiles, instead of means. The full characterization of distributions of posterior quantiles also generalizes the results of Benoît and Dubra (2011), who identify the Bayesian-rationalizable self-ranking data where subjects place themselves relative to the population according to a posterior quantile. Finally, our characterization relies on identifying extreme points of a first-order stochastic dominance interval. Extreme points of orbits under the majorization order (and, hence, of second-order stochastic dominance intervals) have been studied since Hardy, Littlewood, and Pólya (1929), who examine finite-dimensional spaces. Ryff (1967) extends this result to an infinite dimensional space. Kleiner, Moldovanu, and Strack (2021) characterize the extreme points of a subset of orbits under an additional monotonicity assumption, which in turn leads to many economic applications. Extreme points of first-order stochastic dominance intervals exhibit a similar structure—albeit easier to characterize—in the sense that either the stochastic dominance constraints bind on an interval, or there are at most countably many mass points.

In terms of applications, our gerrymandering results are related to the literature on redistricting. Among the closest papers are Owen and Grofman (1988), Friedman and Holden (2008), Gul and Pesendorfer (2010), and Kolotilin and Wolitzky (2020), who adopt the same distribution-based approach and model a district map as a way to split the population distribution of voters. Generally speaking, this work mainly focuses on a political party's optimal gerrymandering when maximizing its expected number of seats. In contrast, our result characterizes the feasible compositions of a legislative body that a district map can induce.

Our second application relates to the Bayesian persuasion literature. Based on the fun-

damental principles outlined by Aumann and Maschler (1995) and Kamenica and Gentzkow (2011), Gentzkow and Kamenica (2016) specialize preferences so that only the posterior means are payoff-relevant. Dworczak and Martini (2019) further generalize the results and provide a characterization of a sender's optimal signals for a general class of mean-based persuasion problems. Kleiner, Moldovanu, and Strack (2021) characterize the extreme points of the feasible set of this convex problem. We complement this literature by providing a foundation for solving persuasion problems where only the posterior quantiles are payoff-relevant. Kolotilin, Corrao, and Wolitzky (2022) consider a general class of persuasion problems and establish sufficient conditions under which particular types of signals are optimal. Their results and our application share a common special case, albeit their results are more general in terms of receiver payoffs, whereas our application is more general in terms of sender payoffs.

In the meantime, the application of market segmentation to a two-sided market has features of both one-sided market segmentation and price discrimination (e.g., Bergemann, Brooks, and Morris 2015; Haghpanah and Siegel 2020, forthcoming; Yang 2022; Elliot, Galeotti, Koh, and Li 2022) as well as matching in a two-sided market (e.g., Hagiu and Jullien 2011; de Cornière 2016; Condorelli and Szentes 2022; Guinsburg and Saraiva 2022).

Finally, the econometric applications are related to problems of inferring the joint distribution from marginals, as studied by Horowitz and Manski (1995) and Cross and Manski (2002); the finance applications are related to the conditional Value-at-Risk measurement of systemic risk introduced by Adrian and Brunnermeier (2016); and the accounting application relates to classification shifting behavior identified in McVay (2006).

Outline The remainder of the paper proceeds as follows. Section 2 establishes the general environment. Section 3 gives the paper's main result. Economic applications follow in Section 4 (gerrymandering), Section 5 (Bayesian persuasion and market segmentation), Section 6 (econometrics), and Section 7 (finance and accounting). Section 8 concludes.

2 Preliminaries

State and Signals Consider a one-dimensional variable $\omega \in \mathbb{R}$. Let $F_0 \in \mathcal{F}$ be the distribution of ω, where $\mathcal{F}$ denotes the collection of distribution functions on $\mathbb{R}$.[1] A particular distribution of interest is the uniform distribution on $[0, 1]$, which is denoted by $U \in \mathcal{F}$. Namely, $U(\omega) = \omega$ for all $\omega \in [0, 1]$. A *signal* for ω is defined as $\mu \in \Delta(\mathcal{F})$ such that

$$\int_{\mathcal{F}} F(\omega)\mu(\mathrm{d}F) = F_0(\omega), \tag{1}$$

[1] $\mathcal{F}$ is endowed with the weak-* topology and the induced Borel σ-algebra, unless otherwise specified.

for all $\omega \in \mathbb{R}$. Let $\mathcal{M}(F_0)$ denote the collection of all signals (under prior distribution F_0). For the ease of exposition, we sometimes write $\mathcal{M}$ instead of $\mathcal{M}(F_0)$ when there is no confusion. From Blackwell's theorem (Blackwell 1953), given any $\mu \in \mathcal{M}$, each $F \in \text{supp}(\mu)$ can be interpreted as a *posterior* for ω obtained via Bayes' rule under a prior F_0, after observing the realization of a signal that is correlated with ω. The marginal distribution of this signal is summarized by μ.

Quantiles and Quantile Selection Rules For any distribution $F \in \mathcal{F}$, let the quantile function F^{-1} be defined as

$$F^{-1}(\tau) := \inf\{\omega \in \mathbb{R} | F(\omega) \geq \tau\},$$

for all $\tau \in (0,1)$.[2] Denote the set of τ-quantiles of F by $\mathbb{Q}^\tau(F) := [F^{-1}(\tau), F^{-1}(\tau^+)]$. Furthermore, we say that a transition probability $r : \mathcal{F} \times [0,1] \to \Delta(\mathbb{R})$ is a *quantile selection rule* if $r(\mathbb{Q}^\tau(F)|F,\tau)) = 1$ for all $F \in \mathcal{F}$ and for all $\tau \in (0,1)$. A quantile selection rule r selects (possibly through randomization) a τ-quantile for every CDF F and for every $\tau \in (0,1)$, whenever it is not unique. Let $\mathcal{R}$ be the collection of all selection rules.

Distributions of Posterior Quantiles For any $\tau \in (0,1)$, for any signal $\mu \in \mathcal{M}$, and for any selection rule $r \in \mathcal{R}$, let $H^\tau(\cdot|\mu, r)$ denote the distribution of the τ-quantile induced by μ and r. That is,

$$H^\tau(\omega|\mu, r) = \int_{\mathcal{F}} r((-\infty, \omega]|F, \tau)\mu(\mathrm{d}F), \tag{2}$$

for all $\omega \in \mathbb{R}$. Our main result characterizes the distributions of posterior quantiles induced by arbitrary signals and selection rules.

Stochastic Dominance Interval Given any $F, F' \in \mathcal{F}$, recall that F *dominates* F' in the sense of first-order stochastic dominance, denoted by $F \succeq F'$ henceforth, if $F(\omega) \leq F'(\omega)$ for all $\omega \in \mathbb{R}$. For any $F, F' \in \mathcal{F}$, such that $F \succeq F'$, denote the set of CDFs that dominate F' and are dominated by F as $\mathcal{I}(F', F)$. That is,

$$\mathcal{I}(F', F) := \{H \in \mathcal{F} | F' \preceq H \preceq F\}.$$

[2]Note that F^{-1} is nondecreasing and left-continuous for all $F \in \mathcal{F}$. Moreover, for any $\tau \in (0,1)$ and for any $\omega \in \mathbb{R}$, $F^{-1}(\tau) \leq \omega$ if and only if $F(\omega) \geq \tau$.

3 Characterization of Distributions of Posterior Quantiles

Our main result characterizes the collection of all possible distributions of posterior quantiles. To state this result, for any $\tau \in (0,1)$, let $\mathcal{H}_\tau$ denote the set of distributions that can be induced by some signal $\mu \in \mathcal{M}$ and selection rule $r \in \mathcal{R}$. Namely,

$$\mathcal{H}_\tau := \{H \in \mathcal{F} | H(\omega) = H^\tau(\omega|\mu, r), \forall \omega \in \mathbb{R}, \text{ for some } \mu \in \mathcal{M}, r \in \mathcal{R}\}.$$

In the meantime, define two distributions $\underline{F}_0^\tau$ and $\overline{F}_0^\tau$ as follows:

$$\underline{F}_0^\tau(\omega) := \min\left\{\frac{1}{\tau}F_0(\omega), 1\right\}, \quad \overline{F}_0^\tau(\omega) := \max\left\{\frac{F_0(\omega) - \tau}{1-\tau}, 0\right\}.$$

Note that $\overline{F}_0^\tau \succeq \underline{F}_0^\tau$ for all $\tau \in (0,1)$. In essence, $\underline{F}_0^\tau$ is the conditional distribution of F_0 in the event that ω is smaller than a τ-quantile of F_0; whereas $\overline{F}_0^\tau$ is the conditional distribution of F_0 in the event that ω is larger than the same τ-quantile. This brings us to our main result.

Theorem 1 (Distributions of Posterior Quantiles)**.** *For any* $\tau \in (0,1)$,

$$\mathcal{H}_\tau = \mathcal{I}(\underline{F}_0^\tau, \overline{F}_0^\tau).$$

Theorem 1 completely characterizes the distributions of posterior τ-quantiles by the stochastic dominance interval $\mathcal{I}(\underline{F}_0^\tau, \overline{F}_0^\tau)$. Figure I illustrates Theorem 1 for the case when $\tau = 1/2$. The distribution $\underline{F}_0^{1/2}$ is colored blue, whereas the distribution $\overline{F}_0^{1/2}$ is colored red. The green dotted curve represents the prior, F_0. According to Theorem 1, any distribution H bounded by $\underline{F}_0^{1/2}$ and $\overline{F}_0^{1/2}$ (for instance, the black curve in the figure) can be induced by a signal $\mu \in \mathcal{M}$ and a select rule $r \in \mathcal{R}$. Conversely, for any signal and for any selection rule, the induced graph of the distribution of posterior τ-quantiles must fall in the area bounded by the blue and red curves.

In what follows, we explain the main steps of the proof of Theorem 1. Details of the proof can be found in the Appendix. To begin with, note that for any signal $\mu \in \mathcal{M}$ and for any $r \in \mathcal{R}$,

$$H^\tau(\omega|\mu, r) \leq \mu(\{F \in \mathcal{F} | F^{-1}(\tau) \leq \omega\}) = \mu(\{F \in \mathcal{F} | F(\omega) \geq \tau\}),$$

for all $\omega \in \mathbb{R}$, where the first inequality holds because the right-hand side corresponds to the distribution of posterior quantiles induced by μ when the lowest τ-quantile is selected with probability 1. Furthermore, for any $\omega \in \mathbb{R}$, if we regard $F(\omega) \in [0,1]$ as a random variable whose distribution is implied by μ, it then follows from (1) that its distribution must be

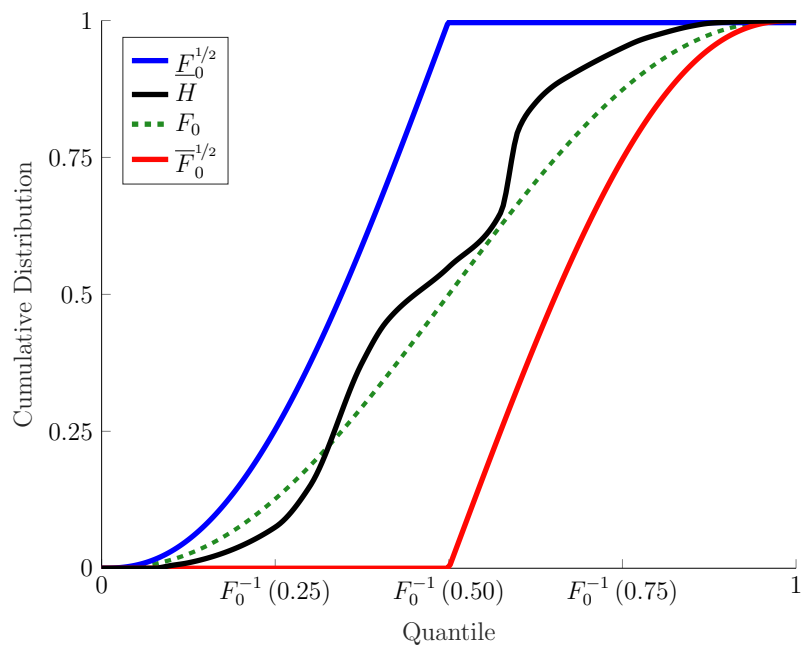


Figure I
Stochastic Dominance Interval

a mean-preserving spread of $F_0(\omega)$. As a result, $\mu(\{F \in \mathcal{F} | F(\omega) \geq \tau\})$ can be at most $\min\{F_0(\omega)/\tau, 1\}$, since otherwise, the mean of $F(\omega)$ can never be $F_0(\omega)$. This implies that $H^\tau(\cdot|\mu, r) \succeq \underline{F}_0^\tau$. A similar argument leads to the conclusion that $H^\tau(\cdot|\mu, r) \preceq \overline{F}_0^\tau$ as well. Thus, $\mathcal{H}_\tau \subseteq \mathcal{I}(\underline{F}_0^\tau, \overline{F}_0^\tau)$.

The converse of the proof is relatively more involved. To begin with, we show that it is without loss to restrict attention to the case where $F_0 = U$. Specifically, for any $\tau \in (0, 1)$ and for any $\omega \in [0, 1]$, let

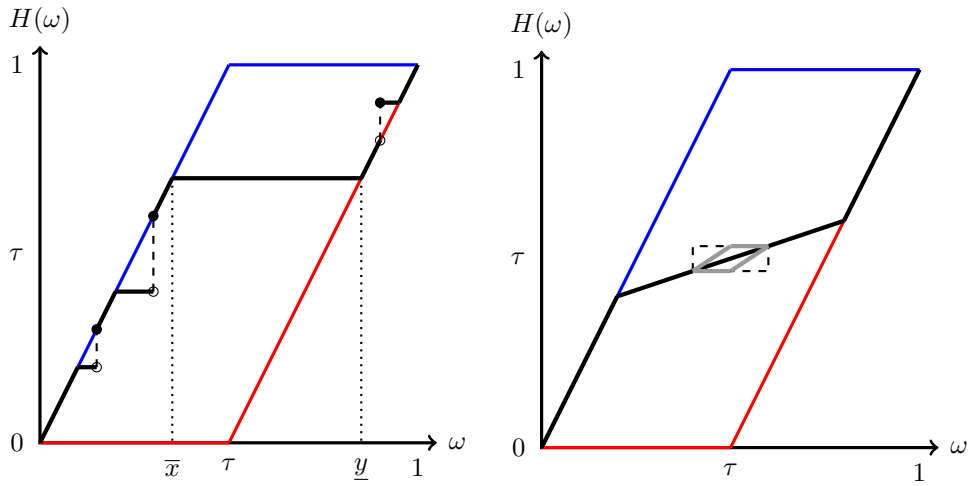
$$\underline{U}^\tau(\omega) := \min\left\{\frac{\omega}{\tau}, 1\right\} \text{ and } \overline{U}^\tau(\omega) := \max\left\{\frac{\omega - \tau}{1 - \tau}, 0\right\},$$

and let $\mathcal{I}_\tau^* := \mathcal{I}(\underline{U}^\tau, \overline{U}^\tau)$. In addition, let $\mathcal{H}_\tau^*$ be the collection of $H \in \mathcal{F}$ such that $H(\omega) = H^\tau(\omega|\tilde{\mu}, \tilde{r})$ for some $\tilde{\mu} \in \mathcal{M}(U)$ and $\tilde{r} \in \mathcal{R}$ for all $\omega \in \mathbb{R}$. Then, we have the following lemma:

Lemma 1. *Consider any $\tau \in (0, 1)$. Then $\mathcal{I}_\tau^* \subseteq \mathcal{H}_\tau^*$ if and only if $\mathcal{I}(\underline{F}_0^\tau, \overline{F}_0^\tau) \subseteq \mathcal{H}_\tau$ for all $F_0 \in \mathcal{F}$.*

By Lemma 1, it suffices to show that $\mathcal{I}_\tau^* \subseteq \mathcal{H}_\tau^*$. To this end, we first characterize the extreme points of $\mathcal{I}_\tau^*$. This characterization is stated in Lemma 2:

Lemma 2. *For any $\tau \in (0, 1)$, H is an extreme point of $\mathcal{I}_\tau^*$ if and only if there exists $0 \leq \underline{x} \leq \overline{x} \leq \tau \leq \underline{y} \leq \overline{y}$; countable sets I, J; and sequences $\{\underline{x}_i, \overline{x}_i\}_{i \in I}, \{\underline{y}_j, \overline{y}_j\}_{j \in J} \subseteq \mathbb{R}$ such*



(A) An Extreme Point **(B) Not an Extreme Point**

Figure II
Extreme Points of $\mathcal{I}_\tau^*$

that $\underline{U}^\tau(\overline{x}) = \overline{U}^\tau(\underline{y})$*; that* $\underline{x} \leq \underline{x}_i \leq \overline{x}_i \leq \underline{x}_{i+1} \leq \overline{x} < \underline{y} \leq \underline{y}_j \leq \overline{y}_j \leq \underline{y}_{j+1} \leq \overline{y}$ *for all* $i \in I, j \in J$*; and that*

$$H(\omega) = \begin{cases} 0, & \text{if } \omega < \underline{x} \\ \underline{U}^\tau(\underline{x}_i), & \text{if } \omega \in [\underline{x}_i, \overline{x}_i) \\ \underline{U}^\tau(\omega), & \text{if } \omega \in [\underline{x}, \overline{x}) \setminus \cup_{i \in I} [\underline{x}_i, \overline{x}_i) \\ \underline{U}^\tau(\overline{x}), & \text{if } \omega \in [\overline{x}, \underline{y}) \\ \overline{U}^\tau(\underline{y}_j), & \text{if } \omega \in [\underline{y}_j, \overline{y}_j) \\ \overline{U}^\tau(\omega), & \text{if } \omega \in [\underline{y}, \overline{y}) \setminus \cup_{j \in J} [\underline{y}_j, \overline{y}_j) \\ 1, & \text{if } \omega \geq \overline{y} \end{cases}, \tag{3}$$

for all $\omega \in \mathbb{R}$*.*

Figure IIA illustrates an extreme point of $\mathcal{I}_\tau^*$. According to Lemma 2, an extreme point H of $\mathcal{I}_\tau^*$ has four cutoffs, $\underline{x}, \overline{x}$ and $\underline{y}, \overline{y}$, such that $\text{supp}(H) \subseteq [\underline{x}, \overline{x}] \cup [\underline{y}, \overline{y}]$. Moreover, on $[\underline{x}, \overline{x}]$, H either coincides with $\underline{U}^\tau$ or is constant over an interval. Similarly, on $[\underline{y}, \overline{y}]$, the left-limit of H either coincides with $\overline{U}^\tau$ or is constant over an interval.

The main idea behind the proof is that, for any $H \in \mathcal{I}_\tau^*$ that does not exhibit this structure—such as the one depicted in Figure IIB—there must exist a rectangle in the interior of the graph of $\mathcal{I}_\tau^*$ such that H is not a step function when restricted to that rectangle. Hence, in that rectangle, H can be split into two distinct nondecreasing functions, as depicted by the gray curves in Figure IIB. This, in turn, implies that H can be split into two distinct nondecreasing functions in $\mathcal{I}_\tau^*$, and hence, H is not an extreme point of $\mathcal{I}_\tau^*$.

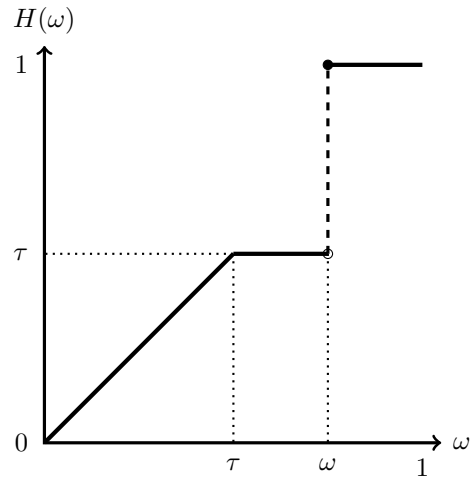


Figure III
An Example of U^ω

Having characterized the extreme points of $\mathcal{I}^*_\tau$, we then show that, for any extreme point H, there exists a signal and a selection rule such that the induced distribution of posterior τ-quantiles coincides with H. Details of the construction can be found in Appendix A.3. The main intuition can be better understood by constructing a signal and a selection rule that attains the boundary $\overline{U}^\tau$. To this end, for any $\omega \in [\tau, 1]$, define a distribution U^ω by:

$$U^\omega(x) := \begin{cases} 0, & \text{if } x < 0 \\ x, & \text{if } x \in [0, \tau) \\ \tau, & \text{if } x \in [\tau, \omega) \\ 1, & \text{if } x \geq \omega \end{cases},$$

for all $x \in \mathbb{R}$. Figure III illustrates U^ω for some $\omega \in [\tau, 1]$. Note that, by construction, $\omega = \max(\mathbb{Q}^\tau(U^\omega))$ for all $\omega \in [\tau, 1]$.

Furthermore, let $\overline{\mu}$ be defined as

$$\overline{\mu}(\{U^\omega | \omega \leq y\}) := \frac{y - \tau}{1 - \tau},$$

for all $y \in [\tau, 1]$. It then follows that $\overline{\mu} \in \mathcal{M}(U)$. Together with the selection rule $\overline{r} \in \mathcal{R}$ that always selects the largest τ-quantile, it follows that $H^\tau(\omega | \overline{\mu}, \overline{r}) = \overline{U}^\tau(\omega)$ for all $\omega \in \mathbb{R}$.[3]

[3]It is noteworthy that many other signals $\mu \in \mathcal{M}(U)$ can also attain the boundary $\overline{U}^\tau$ when properly paired with a selection rule, as long as all states above τ are separated, and each of them is pooled with states below τ so that each is the selected τ-quantile. For instance, when $\tau = 1/2$, the "matching extreme" signal introduced by Friedman and Holden (2008), together with the selection rule that always selects the largest quantile, attains $\overline{U}^{1/2}$ as well. Nonetheless, the "matching extreme" signal cannot attain many other extreme points of $\mathcal{I}^*_{1/2}$. See Section 4.2 for an example.

Finally, since any $\widetilde{H} \in \mathcal{I}_\tau^*$ can be represented as a mixture of the extreme points of $\mathcal{I}_\tau^*$, and since $(\mu, r) \mapsto H^\tau(\cdot|\mu, r)$ is affine, it then follows that $\mathcal{I}_\tau^* \subseteq \mathcal{H}_\tau^*$.[4]

An immediate corollary of Theorem 1 characterizes the set of τ-quantiles of all distributions of posterior $\tilde{\tau}$-quantiles. The corollary can be regarded as the analogue of the law of iterated expectations when means are replaced by quantiles.

Corollary 1 (Law of Iterated Quantiles)**.** *For any $\tau, \tilde{\tau} \in (0,1)$ and for any closed interval $Q \subseteq \mathbb{R}$, $Q = \mathbb{Q}^\tau(H)$ for some $H \in \mathcal{H}_{\tilde{\tau}}$ if and only if $Q \subseteq [(\underline{F}_0^{\tilde{\tau}})^{-1}(\tau), (\overline{F}_0^{\tilde{\tau}})^{-1}(\tau^+)]$.*

Proof. Consider any $H \in \mathcal{H}_{\tilde{\tau}}$. By Theorem 1, $H \in \mathcal{I}(\underline{F}_0^{\tilde{\tau}}, \overline{F}_0^{\tilde{\tau}})$. Thus, $(\underline{F}_0^{\tilde{\tau}})^{-1}(\tau) \leq H^{-1}(\tau) \leq H^{-1}(\tau^+) \leq (\overline{F}_0^{\tilde{\tau}})^{-1}(\tau^+)$, and therefore $\mathbb{Q}^\tau(H) \subseteq [(\underline{F}_0^{\tilde{\tau}})^{-1}(\tau), (\overline{F}_0^{\tilde{\tau}})^{-1}(\tau^+)]$. Conversely, consider any interval $Q = [\underline{\omega}, \overline{\omega}] \subseteq [(\underline{F}_0^{\tilde{\tau}})^{-1}(\tau), (\overline{F}_0^{\tilde{\tau}})^{-1}(\tau^+)]$. Then, let H be defined as

$$H(\omega) := \begin{cases} 0, & \text{if } \omega < \underline{\omega} \\ \tau, & \text{if } \omega \in [\underline{\omega}, \overline{\omega}) \\ 1, & \text{if } \omega \geq \overline{\omega} \end{cases},$$

for all $\omega \in \mathbb{R}$, then $H \in \mathcal{I}(\underline{F}_0^{\tilde{\tau}}, \overline{F}_0^{\tilde{\tau}})$ and $Q = \mathbb{Q}^\tau(H)$. Moreover, by Theorem 1, $H \in \mathcal{H}_{\tilde{\tau}}$, as desired. ■

Another corollary of Theorem 1 characterizes the probability weights assigned to an arbitrary interval $[\underline{\omega}, \overline{\omega}] \subseteq \mathbb{R}$ under a distribution of a posterior τ-quantile.

Corollary 2. *For any $\tau \in (0,1)$ and for any $[\underline{\omega}, \overline{\omega}] \subseteq \mathbb{R}$,*

$$\{H(\overline{\omega}) - H(\underline{\omega}) | H \in \mathcal{H}_\tau\} = [(\overline{F}_0^\tau(\underline{\omega})) - \underline{F}_0^\tau(\overline{\omega}))^+, (\underline{F}_0^\tau(\overline{\omega}) - \overline{F}_0^\tau(\underline{\omega}))].$$

The proof of Corollary 2 is a direct application of Theorem 1 and hence is relegated to the Online Appendix.

Although the characterization of Theorem 1 may seem to rely on selection rules $r \in \mathcal{R}$, the result remains (essentially) the same even when restricted to signals that always induce a unique posterior τ-quantile. Theorem 2 below formalizes this statement. To this end, let $\mathcal{H}_\tau^0$

[4] There are two alternative ways to prove Theorem 1 that bypass the characterization of extreme points. One of them uses a "non-assortative" signal (and thus, not "single-dipped" in the sense of Kolotilin, Corrao, and Wolitzky 2022 and not "matching extreme," in the sense of Friedman and Holden 2008) with binary-support posteriors. Under this signal, any $H \in \mathcal{I}_\tau^*$ can be attained by properly selecting among the multiple posterior quantiles. Another approach is to establish appropriate continuity properties of the mapping $(\mu, r) \mapsto H^\tau(\cdot|\mu, r)$, and then find a proper way to approximate distributions $H \in \mathcal{I}_\tau^*$ by the "rationalizable" data as in Benoît and Dubra (2011). Nonetheless, the proof approach we discussed above applies to characterizing distributions of *unique* posterior quantiles as well (see Theorem 2). Neither of these two alternative approaches can do so, as they rely crucially on the multiplicity of certain posterior quantiles and the selection rule.

be the collection of distributions of posterior τ-quantiles that can be induced by some signal where (almost) all posteriors have a unique τ-quantile. The characterization of $\mathcal{H}_\tau^0$ relates to a family of perturbations of the set $\mathcal{I}(\underline{F}_0^\tau, \overline{F}_0^\tau)$, namely $\{\mathcal{I}(\underline{F}_0^{\tau+\varepsilon}, \overline{F}_0^{\tau-\varepsilon})\}_{\varepsilon>0}$. Note that this family is decreasing in ε under the set inclusion order.

Theorem 2 (Distributions of Unique Posterior Quantiles)**.** *For any $\tau \in (0,1)$, for any $F_0 \in \mathcal{F}$ with full support on an interval in $\mathbb{R}$, there exists $\bar{\varepsilon} > 0$ such that*

$$\bigcup_{0<\varepsilon<\bar{\varepsilon}} \mathcal{I}(\underline{F}_0^{\tau+\varepsilon}, \overline{F}_0^{\tau-\varepsilon}) \subseteq \mathcal{H}_\tau^0 \subseteq \mathcal{I}(\underline{F}_0^\tau, \overline{F}_0^\tau).$$

The proof of Theorem 2 is conceptually similar to the proof of Theorem 1 and thus is relegated to the Online Appendix. The proof relies on a feature of the signals we construct to attain the extreme points of $\mathcal{I}_\tau^*$, where each desired state becomes the unique quantile after a small perturbation. From Theorem 2, even if signals are required to always induce a unique posterior quantile, distributions of posterior τ-quantiles are (almost) characterized by $\mathcal{I}(\underline{F}_0^\tau, \overline{F}_0^\tau)$ as well.[5]

In what follows, we present applications relying on Theorem 1 for the ease of exposition. Nevertheless, all applications have a corresponding version derived from Theorem 2, where either feasible signals are restricted to always induce a unique posterior τ-quantile or selection rules are fixed a priori (e.g., tie-breaking rules in elections are typically by statues).

4 Application I: Gerrymandering

4.1 The Limits of Gerrymandering

With the characterization in hand, our first application is to the consequences of political redistricting. The study of redistricting ranges across many fields: Legal scholars, political scientists, mathematicians, computer scientists, and economists have all contributed to this vast literature.[6] While existing economic theory on redistricting has largely focused on optimal redistricting or fair redistricting mechanisms (e.g., Owen and Grofman 1988; Friedman and Holden 2008; Gul and Pesendorfer 2010; Pegden, Procaccia, and Yu 2017; Ely 2019;

[5]Despite being in a different context and using a different proof approach, Theorem 2 generalizes theorem 1 and theorem 4 of Benoît and Dubra (2011) by allowing for all priors (including non-uniform distributions and those with atoms) and all kinds of partitions (including those with infinitely many elements) of the state space. This, in turn, leads to a characterization of rationalizable data when subjects are asked to forecast their own absolute performance score, rather than their relative position, based on their posterior medians.

[6]See, for example, Shotts (2001); Besley and Preston (2007); Coate and Knight (2007); McCarty, Poole, and Rosenthal (2009); Fryer Jr and Holden (2011); McGhee (2014); Stephanopoulos and McGhee (2015); Alexeev and Mixon (2018).

Friedman and Holden 2020; Kolotilin and Wolitzky 2020), another fundamental question is the scope of redistricting's impact on a legislature. If *any* electoral map can be drawn, what kinds of legislatures can be created? In other words, what are the "limits of gerrymandering"?

Theorem 1 describers the extent to which unrestrained gerrymandering can shape the composition of elected representatives. Specifically, consider an environment in which a continuum of citizens vote, and each citizen has single-peaked preferences over positions on political issues. Citizens have different ideal positions $\omega \in [0, 1]$, and these positions are distributed according to some $F_0 \in \mathcal{F}$.

In this setting, a signal $\mu \in \mathcal{M}$ can be thought of as an electoral *map*, which segments citizens into electoral *districts*, such that a district $F \in \text{supp}(\mu)$ is described by the conditional distribution of the ideal positions of citizens who belong to it. Each district elects a *representative*, and election results at the district-level follow the median voter theorem. That is, given any map $\mu \in \mathcal{M}$, the elected representative of each district F must have an ideal position that is a median of F. When there are multiple medians in a district, the representative's ideal position is determined by a selection rule $r \in \mathcal{R}$.[7]

Given any $\mu \in \mathcal{M}$ and any selection rule $r \in \mathcal{R}$, the induced distribution of posterior medians $H^{1/2}(\cdot|\mu, r)$ can be interpreted as a distribution of the ideal positions of the elected representatives. Meanwhile, the bounds $\underline{F}_0^{1/2}$ and $\overline{F}_0^{1/2}$ can be interpreted as distributions of representatives that only reflect one side of voters' political positions relative to the median of the population. Specifically, $\underline{F}_0^{1/2}$ describes an "all-left" legislature, in which each representative elected has an ideal position that is left of the median voter's ideal. Conversely, $\overline{F}_0^{1/2}$ represents an "all-right" legislature, in which all representatives are positioned to the right of the median voter.[8]

An immediate implication of Theorem 1 is that *any* composition of the legislative body ranging from the "all-left" to the "all-right" can be procured by some map, as summarized by Proposition 1 below. Thus, in the most extreme scenario, unrestrained gerrymandering can lead to a skewed legislature consisting of representatives from only one side.

Proposition 1 (Limits of Gerrymandering)**.** *For any $H \in \mathcal{F}$, the following are equivalent:*

1. *$H \in \mathcal{I}(\underline{F}_0^{1/2}, \overline{F}_0^{1/2})$.*

2. *H is a distribution of the representatives' ideal positions under some map $\mu \in \mathcal{M}$ and some selection rule $r \in \mathcal{R}$.*

[7] Recall that any voting method that meets the Condorcet criterion (e.g., majority voting with two office-seeking candidates) satisfies the median voter property in this setting (Downs 1957; Black 1958).

[8] Gomberg, Pancs, and Sharma (2021) also study how gerrymandering affects the composition of the legislature. However, the authors assume that each district elects a *mean* candidate as opposed to the median.

Knowing the full range of possible legislative compositions allows us to ask a natural question: For a given citizenry distribution F_0, how much can unrestrained gerrymandering contribute to polarization in the legislature? Corollary 2 sheds light on this question by characterizing the share of "moderate" representatives under all maps.

Proposition 2. *For any $\beta \in [1/2, 1]$ and for any $\eta \in [0, 1]$, the following are equivalent:*

1. $\max\{(4\beta - 3), 0\} \leq \eta \leq \min\{4\beta - 1, 1\}$.

2. *There exists a map and a selection rule such that the share of representatives with ideal positions in $[F_0^{-1}(1-\beta), F_0^{-1}(\beta^+)]$ is η.*

Proof. For any $\beta \in [1/2, 1]$, $(\overline{F}_0^{1/2}(F_0^{-1}(\beta^+)) - \underline{F}_0^{1/2}(F_0^{-1}(1-\beta)))^+ = \max\{4\beta - 3, 0\}$, whereas $\underline{F}_0^{1/2}(F_0^{-1}(\beta^+)) - \overline{F}_0^{1/2}(F_0^{-1}(1-\beta)) = \min\{4\beta - 1, 1\}$. The result then immediately follows from Corollary 2. ■

For any $\beta \in [1/2, 1]$, Proposition 2 implies that gerrymandering can lead to a legislature where the share of "β-moderates" (i.e., those with ideal positions within $[F_0^{-1}(1-\beta), F_0^{-1}(\beta^+)]$) is as small as $(4\beta - 3)^+$. In particular, unrestrained gerrymandering can lead to a legislature as polarized as having no representatives with positions in the interquartile range of F_0. From this perspective, in addition to the skewed "all-left" and "all-right" legislatures, many other extreme compositions are possible under unrestrained gerrymandering, including a polarized chamber with no elected moderates.

Having a complete characterization of compositions of the legislature that can arise under gerrymandering, we may further explore the set of possible *legislative outcomes* that can be enacted. To this end, we must impose further assumptions on the congressional voting method that enacts legislation. One example is that enacted legislation must be a median of the representatives.[9] Under this system, an immediate implication of Corollary 1 is that the set of achievable legislative outcomes coincides with the interquartile range of the citizenry's ideal positions $[F_0^{-1}(1/4), F_0^{-1}(3/4^+)]$.

More generally, we may regard a legislative voting method as a mapping from the distribution of representatives' ideal positions to a legislative outcome. One natural requirement for such mappings is that they reflect the will of a majority whenever that will is unambiguous. In other words, for any distribution of representatives' ideal positions with more than $1/2$ of the representatives having the same ideal position $\omega \in [0, 1]$, the legislative voting procedure must yield outcome ω. Notice that a voting system that always enacts a median position of the representatives satisfies this requirement. As shown by Proposition 3 below, regardless of

[9] See McCarty, Poole, and Rosenthal 2001; Bradbury and Crain 2005; and Krehbiel 2010 for evidence that the median legislator is decisive. See also Cho and Duggan (2009) for a microfoundation.

the exact voting procedure adopted, the set of possible legislative outcomes always includes the interquartile range.

Proposition 3. *Consider any function* $\mathbb{C} : \mathcal{H}_{1/2} \to [0,1]$. *Suppose that* $\mathbb{C}(H) = \omega$ *for all* H *that assigns probability greater than* $1/2$ *to* ω. *Then,* $\mathbb{C}(\mathcal{H}_{1/2}) \supseteq (F_0^{-1}(1/4), F_0^{-1}(3/4^+))$.

Proof. Consider any $\omega \in (F^{-1}(1/4), F^{-1}(3/4^+))$. Let H_ω be a distribution that assigns probability $2 \cdot \min\{F_0(\omega), 1 - F_0(\omega)\}$ to ω, and probability $1 - 2 \cdot \min\{F_0(\omega), 1 - F_0(\omega)\}$ to $F_0^{-1}(1/2)$. Then $H_\omega \in \mathcal{I}(\underline{F}_0^{1/2}, \overline{F}_0^{1/2})$. Therefore, by Theorem 1, $H_\omega \in \mathcal{H}_{1/2}$, which, in turn, implies that $\mathbb{C}(H_\omega) = \omega$, as desired. ■

Proposition 3 shows that, under a wide variety of legislative voting procedures, unrestrained gerrymandering can lead to *any* legislative outcome within the interquartile range of citizens' ideal positions, even if all district-level elections adhere to the median voter property, and only the population medians are the *Condorcet winners* in this setting.[10] Moreover, since the interquartile range expands under a more polarized distribution of political views, Proposition 3 also implies that unrestrained gerrymandering can lead to more extreme legislation during more polarized times.

Knowing that the set of possible legislative outcomes far exceeds the Condorcet winners, we next ask: Which legislative outcomes can *defeat* the Condercet winners by securing a majority of support among representatives elected under some map? Proposition 4 below characterizes the set of legislative outcomes that are preferred by a fraction $\alpha \in [1/2, 1]$ of the representatives over any population medians under some map. To state this result, we let $\underline{\omega}(\alpha) := \max\{2F_0^{-1}(\alpha/2) - F_0^{-1}(1/2), 0\}$ and $\overline{\omega}(\alpha) := \min\{2F_0^{-1}(1 - \alpha/2) - F_0^{-1}(1/2), 1\}$.

Proposition 4. *For any* $\omega \in [0,1]$ *and for any* $\alpha \in [1/2, 1]$, *the following are equivalent:*

1. $\omega \in [\underline{\omega}(\alpha), \overline{\omega}(\alpha)]$.
2. *There exists a map and a selection rule such that* ω *is preferred to any population median by at least* α *share of the representatives.*

Proof. See Appendix A.4. ■

According to Proposition 4, even though the population medians are Condorcet winners, for any legislative outcome ω in $[\underline{\omega}(\alpha), \overline{\omega}(\alpha)]$, there exists a gerrymandered map that would secure ω with α-absolute majority of support among representatives. A special case for

[10] Recall that a Condorcet winner is defined as an outcome that has majority support when compared to any other alternative. As every citizen has single-peaked preferences over positions in $[0,1]$, a Condorcet winner always exists, and the set of Condorcet winners coincides with the population medians $\mathbb{Q}^{1/2}(F_0)$.

this result is when F_0 has a symmetric and quasi-convex density. In this case, $\underline{\omega}(1/2) = 0$ and $\overline{\omega}(1/2) = 1$. That is, under a polarized population distribution (even only slightly), *any* outcome in $[0, 1]$ can defeat the population medians by simple majority rule under some map, which is arguably a complete reversal of the population medians' Condorcet property. In the meantime, note that $\overline{\omega}$ is decreasing in α and $\underline{\omega}$ is increasing in α. Moreover, $\overline{\omega}(1) = F_0^{-1}(1/2^+)$ and $\underline{\omega}(1) = F_0^{-1}(1/2)$. This suggests that raising the voting threshold for an alternative to beat the population median, such as requiring an absolute majority or unanimous support, can mitigate the impact of gerrymandering in this regard.

Remark 1 (Districts on a Geographic Map). In practice, election districts are drawn on a geographic map. Drawing districts in this manner can be regarded as partitioning a two-dimensional space that is spanned by latitude and longitude. More specifically, let a convex and compact set $\Theta \subseteq [0, 1]^2$ denote a geographic map. Suppose that every citizen who resides at the same location $\theta \in \Theta$ shares the same ideal position $\boldsymbol{\omega}(\theta)$, where $\boldsymbol{\omega} : \Theta \to [0, 1]$ is a measurable function. Furthermore, suppose that citizens are distributed on Θ according to a density function $\phi > 0$. Under this setting, theorem 1 of Yang (2020) ensures that for any $\mu \in \mathcal{M}$ with a countable support, there exists a countable partition of Θ, such that the distributions of citizens' ideal positions within each element coincide with the distributions in the support of μ. If we further assume that $\boldsymbol{\omega}$ is non-degenerate, in the sense that each of its indifference curves $\{\theta \in \Theta | \boldsymbol{\omega}(\theta) = \omega\}_{\omega \in [0,1]}$ is isomorphic to the unit interval, then theorem 2 of Yang (2020) ensures that for any $\mu \in \mathcal{M}$, there exists a partition on Θ that generates the same distributions in each district. Therefore, the splitting of the distribution of citizens' ideal positions has an exact analogue to the splitting of geographic areas on a physical map.

4.2 Optimal Gerrymandering with Aggregate Uncertainty

In addition to characterizing the set of possible compositions of a legislative body that all maps can induce, Theorem 1 also sheds lights on optimal gerrymandering problems in the presence of aggregate uncertainty. Consider the same unit mass of citizens whose ideal positions ω are distributed according to F_0. Suppose that there is a map drawer who designs the map of districts. The map drawer's objective depends on two political parties' seat shares in the legislative body. Specifically, consider a model with aggregate uncertainty as in Friedman and Holden (2008) and Kolotilin and Wolitzky (2020), but abstract away individual uncertainty. Suppose that $X \sim G \in \mathcal{F}$ is an aggregate shock to citizens' views on the two parties. Given any realization $x \in [0, 1]$, a citizen with ideal position $\omega \in [0, 1]$ votes for one party (Party 1) if and only if $\omega \geq x$; whereas, she votes for the other party (Party 0) if and only if $\omega < x$. Given a map $\mu \in \mathcal{M}$, a party wins a district if more than 50% of the citizens

in that district vote for that party. Ties are broken by a tie-breaking rule $r \in \mathcal{R}$ selected by the map drawer.

The map drawer chooses a map $\mu \in \mathcal{M}$ and a tie-breaking rule $r \in \mathcal{R}$ to maximize her payoff $W(s)$, where $W : [0,1] \to \mathbb{R}$ and s denotes the share of districts that Party 1 wins (i.e., Party 1's seat share of the legislative body). Note that, for any district $F \in \mathcal{F}$ and for any realized aggregate shock $x \in [0,1]$, Party 1 wins the district if $x \leq F^{-1}(1/2)$. Therefore, given any realized x, the map drawer's payoff under map μ and tie-breaking rule r is $1 - H^{1/2}(x^-|\mu, r)$, and hence, the map drawer's problem can be written as

$$\sup_{\mu \in \mathcal{M}, r \in \mathcal{R}} \int_0^1 W\left(1 - H^{1/2}(x^-|\mu, r)\right) G(\mathrm{d}x).$$

But this problem, by Theorem 1, is equivalent to

$$\sup_{H \in \mathcal{I}(\underline{F}_0^{1/2}, \overline{F}_0^{1/2})} \int_0^1 W(1 - H(x^-))G(\mathrm{d}x). \tag{4}$$

If W is increasing, then the solution of (4) is $\overline{F}_0^{1/2}$. This coincides with the solution stated in proposition 3 of Kolotilin and Wolitzky (2020), which, in turn, is the full-information limit of the solution in Friedman and Holden (2008). But more broadly, Theorem 1 provides solutions to the map drawer's problem for *any* objective function W, not necessarily increasing.[11] For instance, the map drawer might be a bipartisan commission who wants to maintain a balanced seat share, so that W is upper-semicontinuous, symmetric and quasi-concave with a peak at $s = 1/2$. This, in turn, implies that the solution is $H^* \in \mathcal{I}(\underline{F}_0^{1/2}, \overline{F}_0^{1/2})$, where

$$H^*(\omega) := \begin{cases} \underline{F}_0^{1/2}(\omega), & \text{if } \omega < F_0^{-1}(1/4) \\ \frac{1}{2}, & \text{if } \omega \in [F_0^{-1}(1/4), F_0^{-1}(3/4^+)) \\ \overline{F}_0^{1/2}(\omega), & \text{if } \omega \geq F_0^{-1}(3/4^+) \end{cases}, \tag{5}$$

for all $\omega \in \mathbb{R}$.[12] For detailed arguments, see the Online Appendix. More generally, note that

[11] The main result of Kolotilin and Wolitzky (2020) generalizes their proposition 3 by introducing individual shocks on top of the aggregate shock. With individual shocks, the ex-post seat shares are no longer functions of district medians, and thus Theorem 1 does not apply. That being said, their main result maintains the assumption that W is increasing; whereas, Theorem 1 applies to any functional form of W. In this regard, Theorem 1 relates to their main result by generalizing their proposition 3 as well, but along a different dimension.

[12] Note that the solution H^*—which is also the (essentially) unique solution of (4) when both F_0 and G have full support on a common interval and when W is strictly quasi-concave—assigns zero probability to the interval $[F_0^{-1}(1/4), F_0^{-1}(3/4)]$, and hence, the "matching extreme" map by Friedman and Holden (2008) is not optimal regardless of the selection rule in this case.

with a change of variable, (4) can be written as

$$\sup_{\widetilde{H}\in\mathcal{I}(G\circ(\overline{F}_0^{1/2})^{-1},G\circ(\underline{F}_0^{1/2})^{-1})}\int_0^1 W(1-s)\widetilde{H}(\mathrm{d}s),$$

which becomes a linear programming problem. In particular, the extreme points of the feasible set can readily be characterized using similar arguments as in the proof of Lemma 2.

5 Application II: Bayesian Persuasion and Market Segmentation

Here, we apply our characterization to topics in Bayesian persuasion, cheap talk, and consumer market segmentation. More detailed arguments for this section can be found in the Online Appendix.

5.1 Quantile-Based Bayesian Persuasion

Consider the canonical Bayesian persuasion problem of Kamenica and Gentzkow (2011). A state $\omega \in \mathbb{R}$ is distributed according to a common prior F_0. A sender chooses a signal $\mu \in \mathcal{M}$ to inform the receiver, who then picks an action $a \in A$ after seeing the signal's realization. The ex-post payoffs of the sender and receiver are $u_S(\omega, a)$ and $u_R(\omega, a)$, respectively. Kamenica and Gentzkow (2011) show that the sender's optimal signal and the value of persuasion can be characterized by the concave closure of the function $\hat{v} : \mathcal{F} \to \mathbb{R}$, where $\hat{v}(F) := \mathbb{E}_F[u_S(\omega, a^*(F))]$ and $a^*(F) \in A$ is the sender-preferred optimal action under posterior $F \in \mathcal{F}$.

When $|\mathrm{supp}(F_0)| \geq 2$, this "concavafication" method requires finding the concave closure of a multi-variate function, which is known to be computationally challenging, especially when $|\mathrm{supp}(F_0)| = \infty$. For tractability, many papers have restricted attention to preferences where the only payoff-relevant statistic for the sender is the posterior mean (i.e., $\hat{v}(F)$ is measurable with respect to $\mathbb{E}_F[\omega]$). See, for example, Gentzkow and Kamenica (2016); Kolotilin, Li, Mylovanov, and Zapechelnyuk (2017); Kolotilin (2018); Dworczak and Martini (2019); Ali, Haghpanah, Lin, and Siegel (2022); and Kolotilin, Mylovanov, and Zapechelnyuk (forthcoming). A natural analogue of this "mean-based" setting is for the payoffs to depend only on the posterior quantiles. Our main characterization provides means to solve this class of problems defined by "quantile-based" payoffs.

Specifically, suppose that the sender's and receiver's payoffs are such that there exists

$\tau \in (0,1)$ and a measurable function $v_S : \mathbb{R} \to \mathbb{R}$ in which

$$\hat{v}(F) = \sup_{\omega \in \mathbb{Q}^\tau(F)} v_S(\omega), \tag{6}$$

for all $F \in \mathcal{F}$. Under this assumption, Theorem 1 implies that the sender's problem can be rewritten as

$$\sup_{H \in \mathcal{I}(\underline{F}_0^\tau, \overline{F}_0^\tau)} \int_{\mathbb{R}} v_S(\omega) H(\mathrm{d}\omega). \tag{7}$$

Namely, the sender simply selects a distribution belonging to the stochastic dominance interval $\mathcal{I}(\underline{F}_0^\tau, \overline{F}_0^\tau)$ that maximizes the expected value of $v_S(\omega)$, rather than concavafying the infinite-dimensional functional $\hat{v}$. A significant benefit of this simplification is that the sender only needs to solve a constrained maximization problem with an affine objective and a well-behaved feasible set.[13] In what follows, we demonstrate this simplification by revisiting the two examples of Kamenica and Gentzkow (2011).

Lobbying under the Absolute Loss Function

A politician (receiver) chooses a one dimensional policy $a \in \mathbb{R}$ to match the state $\omega \in \mathbb{R}$, which is unknown to the politician and follows a common prior $F_0 \in \mathcal{F}$. The lobbyist (sender) can choose any signal for ω to affect the politician's choice of policy. Kamenica and Gentzkow (2011) assume that the lobbyist's payoff is $u_S(\omega, a) = -(a - \alpha\omega - (1-\alpha)\omega_0)^2$, for some fixed $\omega_0 \in \mathbb{R}$ and $\alpha \in [0,1]$; and the politician's payoff is $u_R(\omega, a) = -(a-\omega)^2$. The quadratic loss structure simplifies the lobbyist's problem into a mean-based persuasion problem that can easily be solved analytically.

Of course, one may argue that the quadratic loss structure is specific, and the general lobbying problem remains difficult to solve. Nonetheless, Theorem 1 lets one solve another parameterization of this problem. Instead of quadratic loss, suppose now that the politician's payoff is given by the *absolute loss*: $u_R(\omega, a) = -|a - \omega|$. Also, suppose that the lobbyist's payoff is state-independent: $u_S(\cdot, a)$ is constant for all a. Then, for any posterior $F \in \mathcal{F}$, the politician's optimal actions are $\mathbb{Q}^{1/2}(F)$, and thus, the lobbyist's problem can be written as:

$$\sup_{H \in \mathcal{I}(\underline{F}_0^{1/2}, \overline{F}_0^{1/2})} \int_{\mathbb{R}} v_S(\omega) H(\mathrm{d}\omega), \tag{8}$$

which can now be solved analytically. For instance, suppose that v_S is increasing (resp.,

[13] Indeed, notice that $\mathcal{I}(\underline{F}_0^\tau, \overline{F}_0^\tau)$ is compact under the weak-* topology, is a complete lattice under the partial order $\preceq$, and is a convex subset of a linear space whose extreme points can readily be derived from Lemma 1 and Lemma 2.

decreasing). Then the lobbyist's optimal signal is $\overline{F}_0^{1/2}$ (resp., $\underline{F}_0^{1/2}$). Alternatively, suppose that $\Omega = [0,1]$ and that v_S is upper-semicontinuous and quasi-concave with peak at $\omega_0 \in [0, F_0^{-1}(\tau)]$.[14] Let $H^{**} \in \mathcal{I}(\underline{F}_0^{1/2}, \overline{F}_0^{1/2})$ be defined as:

$$H^{**}(\omega) := \begin{cases} 0, & \text{if } \omega < \omega_0 \\ \underline{F}_0^{1/2}(\omega), & \text{if } \omega \geq \omega_0 \end{cases}.$$

Then H^{**} is a solution to the lobbyist's problem. In general, for any (measurable) function $v_S : \mathbb{R} \to \mathbb{R}$, note that, by the same arguments as in the proof of Lemma 1, the lobbyist's problem can be characterized by solving

$$\sup_{H \in \mathcal{I}^*_{1/2}} \int_{\mathbb{R}} v_S(\omega) H(\mathrm{d}\omega).$$

Thus, by Lemma 1 and Lemma 2, it suffices to search for the optimal signal within the class of distributions satisfying (3).

Supplying Product Information

The second example from Kamenica and Gentzkow (2011) has a seller choosing the information structure of its product when faced with a single potential buyer. Consider the following general framework for this environment: There is one seller and one buyer; the seller has unit supply, whereas the buyer has unit demand. The product has characteristic $\theta \in \Theta$ that is unobservable to the buyer, and the buyer has a private type $x \in X$. Both θ and x are independently drawn from a common prior. The interpretation of θ and x is that θ describes the product's features, whereas x is the buyer's personal taste. Given price $p \geq 0$, characteristic $\theta \in \Theta$, and type $x \in X$, the buyer's indirect utility of having the product is $u(\theta, x, p) \in \mathbb{R}$. The buyer can choose whether to buy the product or retain an outside option worth $\bar{u} \in \mathbb{R}$. The seller chooses a price p and a signal that informs the buyer about θ so as to maximize revenue.

Kamenica and Gentzkow (2011) consider a special case in which the buyer is an Expected Utility maximizer, has quasi-linear preferences and no private type, and $\Theta = \mathbb{R}_+$, $\bar{u} = 0$, $u(\theta, x, p) = \theta - p$. They then conclude that an optimal signal for the seller, given price p, is to induce at most two signal realizations—one "low", one "high"—such that the buyer's posterior expected gains from trade under the "high" signal realization equals $\max\{p, \mathbb{E}[\theta]\}$. When the seller also optimizes with respect to price, it then follows that the seller's optimal

[14] In the Online Appendix, we also characterize the solutions when v_S is quasi-convex, which have the same feature as H^* defined in (5) qualitatively.

price equals the expected gains from trade and the optimal signal reveals no information, as this allows the seller to fully extract the surplus.[15]

Our main result enables us to explore different specifications of the general model above and extend the results in Kamenica and Gentzkow (2011) by allowing for private tastes. Suppose that $\Theta = X = [0,1]$, $\bar{u} = 0$, the buyer is an Expected Utility maximizer, and $u(\theta, x, p) = v \cdot \mathbf{1}\{\theta \geq x\} - p$ for all θ, x, and p, for some $v \geq 0$. Let F_0 and G denote the priors from which θ and x are drawn, respectively. An interpretation of this payoff structure is that the buyer owns a complement of the seller's product that is only compatible with certain characteristics. For example, the buyer owns a hardware of generation $x \in [0,1]$ that is only compatible with software of generation $\theta \geq x$.

Given any posterior F for θ, the buyer would purchase the product whenever $v(1 - F(x^-)) - p \geq 0$, which in turn is equivalent to $F^{-1}((1 - p/v)^+) \geq x$. As a result, given price $p \geq 0$, the seller's optimal signal can be characterized by the solution of

$$\sup_{H \in \mathcal{I}(\underline{F}_0^{(1-p/v)}, \overline{F}_0^{(1-p/v)})} \int_0^1 (1 - H(x^-))G(\mathrm{d}x),$$

and hence, the optimal revenue given p is

$$p \int_0^1 (1 - \overline{F}_0^{(1-v/p)}(x))G(\mathrm{d}x). \tag{9}$$

The optimal price then maximizes (9). It is noteworthy that, unlike the model without private types, the seller does not fully extract the surplus and discloses some information about θ to the buyer.[16]

Another specification of the general model above is to suppose that the buyer has Quantile-Maximizing preferences, rather than maximizing expected utility. This model of preferences was developed in Manski (1988), Chambers (2007), Rostek (2010), and de Castro and Galvao (2021). When selecting among lotteries, a quantile-maximizing individual chooses the one that gives the highest quantile of the utility distribution. For example, the person might

[15] A version of this problem with consumer search or constant marginal cost is also studied by Anderson and Renault (2006).

[16] This specification essentially coincides with example 3 of Kolotilin, Corrao, and Wolitzky (2022), which, in turn, is equivalent to a persuasion model where both the state and the receiver's action are in $[0,1]$ and $u_R(\omega, a) = \min\{a, \omega\} - \kappa a$, for some $\kappa \in [0,1]$. In the meantime, the sender has an increasing, state-independent payoff. We thank Alexander Wolitzky for pointing out this connection. While the main interest of Kolotilin, Corrao, and Wolitzky (2022) is to explore the qualitative features of optimal signals under more general receiver payoffs, our main result complements theirs since the characterization of Theorem 1 allows us to generalize the sender's payoff to non-monotone functions of actions, in which case "single-dipped" signals are not optimal in general, as demonstrated in Section 4.2 and in the Online Appendix.

maximize median utility instead of mean utility, as she would if she exhibited Expected Utility preferences.

Specifically, suppose that $u(\theta, x, p) = \tilde{u}(\theta, p)$ for some $\tilde{u} : \Theta \times \mathbb{R}_+ \to \mathbb{R}$ so that the buyer has no taste differentiation. Moreover, suppose that the buyer is a $\tau \in (0,1)$-quantile maximizer, as defined by Rostek (2010). Under these assumptions, note that it is without loss to represent a signal for θ by $\mu \in \mathcal{M}(F_p)$, where F_p denotes the distribution of the buyer's utility given price p. Having Quantile-Maximizing preferences, the buyer will purchase the product under a posterior $F \in \mathcal{F}$ if $F^{-1}(\tau^+) > \bar{u}$ and will not purchase if $F^{-1}(\tau) < \bar{u}$, for some $\tau \in (0,1)$. By Theorem 1, the seller's problem can be written as

$$\sup_{H \in \mathcal{I}(\underline{F}_p^\tau, \overline{F}_p^\tau)} p(1 - H^\tau(\bar{u}^-)),$$

and hence, the seller's optimal revenue given price p is $p(1 - \overline{F}_p^\tau(\bar{u}^-))$. The seller's optimal price can then be found by choosing p to maximize $p(1 - \overline{F}_p^\tau(\bar{u}^-))$. Here too, the seller optimally discloses some information.[17]

Quantile-Based Cheap Talk with Transparent Motives

Another setting related to quantile-based persuasion problems is the class of quantile-based cheap talk games with transparent motives. Consider the setting of Lipnowski and Ravid (2020): A sender observes the state ω and can send a message to the receiver. The receiver observes the message and then takes an action. The sender does not have commitment power and has a state-independent payoff. Suppose further that the state ω and the action a are both in $\mathbb{R}$. Theorem 2 of Lipnowski and Ravid (2020) characterizes the sender's best equilibrium payoffs by the quasi-concave envelope of $\hat{v} : \mathcal{F} \to \mathbb{R}$. Just as when the sender has commitment, computing this quasi-concave envelope can be challenging. Sharper analytical solutions would require further specifications of the payoffs.

Theorem 1 allows us to further characterize the sender's equilibrium payoffs under the assumption that payoffs are "quantile-based" as defined in (6).[18] Let the prior distribution of the state ω be F_0. Assume that $\text{supp}(F_0) = [0,1]$ and that v_S is upper-semicontinuous. Note that for any $v^* \in \mathbb{R}$, the set $\{\omega \in [0,1] | v_S(\omega) < v^*\}$ is open and hence can be written as the union of countably many disjoint open intervals $\{(\underline{\omega}^i_{v^*}, \overline{\omega}^i_{v^*})\}_{i=1}^{I_{v^*}}$ for some $I_{v^*} \leq \infty$.

[17] In practice, this optimal signal has the seller sharing information with the buyer to distinguish products, conditional on the products having high values, but also retaining information to prevent the buyer from completely discerning high-valued products from low-valued ones. For instance, the seller might disclose that a handbag were made in Italy, but withhold the means of production (machine versus handmade).

[18] This can be achieved by assuming that the receiver's optimal actions given posterior F are τ-quantiles of F (e.g., $u_R(\omega, a) = -|\omega - a|$).

Theorem 1 of Lipnowski and Ravid (2020) implies that $v^* \geq v_S(F_0^{-1}(\tau))$ is an equilibrium payoff for the sender if and only if there exists a signal $\mu \in \mathcal{M}$ and a selection $r \in \mathcal{R}$ such that, for $\omega \sim H^\tau(\cdot|\mu, r)$, one has $v_S(\omega) \geq v^*$ with probability 1. By Theorem 1, this condition in turn is equivalent to the following: $v^* \geq v_S(F_0^{-1}(\tau))$ is an equilibrium payoff for the sender if and only if there exists $H \in \mathcal{I}(\underline{F}_0^\tau, \overline{F}_0^\tau)$ such that H is constant on $(\underline{\omega}_{v^*}^i, \overline{\omega}_{v^*}^i)$ for all i. By Corollary 2, it follows that the sender's equilibrium payoffs that are higher than the babbling equilibrium payoff can be characterized by the set of $v^* \geq v_S(F_0^{-1}(\tau))$ such that $\overline{F}_0^\tau(\overline{\omega}_{v^*}^i) \leq \underline{F}_0^\tau(\underline{\omega}_{v^*}^i)$ for all i.

For example, suppose that v_S is (strictly) quasi-concave and F_0 has full support. Without loss of generality, suppose that the maximum of v_S is attained at $\omega_0 \geq F_0^{-1}(\tau) > 0$. Then, for any $v^* > v_S(F_0^{-1}(\tau))$, there exists a unique interval $[\underline{\omega}, \overline{\omega}]$ with $F_0^{-1}(\tau) < \underline{\omega}$, such that $v_S(\omega) \geq v^*$ if and only if $\omega \in [\underline{\omega}, \overline{\omega}]$. Since $\overline{F}_0^\tau(\underline{\omega}) > 0 = \underline{F}_0^\tau(0)$, Corollary 2 implies that there does not exist any $H \in \mathcal{I}(\underline{F}_0^\tau, \overline{F}_0^\tau)$ that is constant on $[0, \underline{\omega}]$. Thus, the sender cannot achieve any payoff higher than the babbling equilibrium payoff $v_S(F_0^{-1}(\tau))$.

Nonetheless, the sender's equilibrium payoff is *not* characterized by the quasi-concave envelope of v_S. (It is, however, characterized by the quasi-concave closure of $\hat{v}$ according to theorem 2 of Lipnowski and Ravid 2020.) To see this, suppose that $\tau = 1/2$, that $v_S(\omega) = (\omega - 1/2)^2$, and that $F_0 = U$. In this setting, the quasi-concave envelope of v_S is a constant $1/4$. However, for any $v^* \geq v_S(1/2)$, $v_S(\omega) \leq v^*$ if and only if $\omega \in [1/2 - \sqrt{v^*}, 1/2 + \sqrt{v^*}]$. Thus, the highest equilibrium payoff for the sender is v^* such that $\sqrt{v^*} = 1/4$, which equals $1/16$ and is smaller than $1/4$. A more detailed argument can be found in the Online Appendix.

5.2 Optimal Market Segmentation with a Fixed Thickness Constraint

In addition to being a signal in Blackwell's sense, $\mu \in \mathcal{M}$ has another common interpretation as a *market segmentation* that splits a single market into several segments to facilitate price discrimination (see, for instance, Bergemann, Brooks, and Morris 2015; Haghpanah and Siegel 2020; Yang 2022; Haghpanah and Siegel forthcoming; Elliot, Galeotti, Koh, and Li 2022). From this perspective, Theorem 1 enables an exploration of optimal market segmentation in a two-sided market. When the market is two-sided, market segmentation involves another dimension; namely, one needs to describe how segments on one side are matched with those on the other. (See, for instance, Hagiu and Jullien 2011; de Cornière 2016; Condorelli and Szentes 2022; Guinsburg and Saraiva 2022.)

Consider a two-sided market (e.g., ride sharing) for an object (e.g., a car ride). The demand side is populated by a unit mass of agents (riders) who have unit demands for rides. Their values ω for a ride are distributed according to a distribution $F_0 \in \mathcal{F}$. The supply side is populated by a mass $\tau \in (0, 1)$ of agents (drivers) who have unit supply. Total

supply is inelastic (e.g., during peak hours at a major airport). A third-party platform (a ride-sharing app) can segment both sides of the market to affect prices, which are, in turn, determined by the market-clearing condition in each segment. Specifically, $\mu \in \mathcal{M}(F_0)$ can be regarded as a demand-side market segmentation. A segmentation μ induces market segments $\{F | F \in \text{supp}(\mu)\}$, and each segment is described by the distribution of riders' values within that segment.

For many practical reasons (e.g., regulation, fairness, corporate image, match efficiency, or customer satisfaction), platforms can rarely segment both sides of the market arbitrarily. Instead, they typically face several constraints when segmenting the market. One practical constraint is that the *market thickness* must be held fixed across all market segments. In this setting, where the supply is perfectly inelastic, a market thickness constraint means that the sizes of segments on the demand side must completely determine the sizes of segments on the supply side, so that the ratio of supply and demand remains τ in each segment.[19]

In this setting, Theorem 1 provides a way for observers (e.g., regulators or econometricians) to verify whether the segmentation created by the platform adheres to the thickness constraint. Verifying compliance with a fixed thickness constraint could be challenging, as it requires full knowledge of how the market is segmented, which might be difficult to obtain. Nonetheless, Theorem 1 implies that it is sufficient to observe the distribution of prices across segments to check compliance. To see this, note that for any segment $F \in \mathcal{F}$, given that the ratio of supply and demand is τ, the implied market-clearing price for this segment must be in $\mathbb{Q}^{(1-\tau)}(F)$. (Notice that the function $\tau \mapsto F^{-1}(1-\tau)$ can be regarded as the *inverse demand* of segment F.) As a result, a price distribution is consistent with the thickness constraint only if it falls in $\mathcal{I}(\underline{F}_0^{(1-\tau)}, \overline{F}_0^{(1-\tau)})$.

Furthermore, Theorem 1 provides a characterization of all possible outcomes that can be induced by a market segmentation satisfying a fixed thickness constraint, which we present in the next proposition.

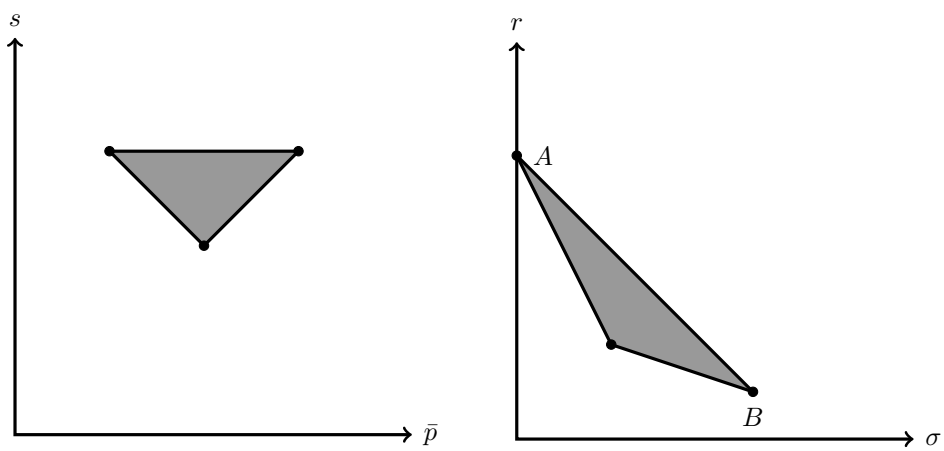
Proposition 5. $(\bar{p}, s) \in \mathbb{R}^2$ *is a pair of average price and total surplus under some market segmentation with fixed thickness* $\tau \in (0,1)$ *if and only if*

$$\tau \int_0^1 F_0^{-1}(1-x)\,\mathrm{d}x \leq s \leq \int_0^\tau F_0^{-1}(1-x)\,\mathrm{d}x \tag{10}$$

and

$$\frac{\int_0^1 F_0^{-1}(1-x)\,\mathrm{d}x - s}{1-\tau} \leq \bar{p} \leq \frac{s}{\tau}. \tag{11}$$

[19] Practically speaking, a thickness constraint would imply that all riders would have to wait approximately the same time before being matched with a driver.



(A) Price-Surplus Pair **(B) Surplus Division**

Figure IV
Two-Sided Market Possible Outcomes

Proof. See Appendix A.5 ■

Figure IVA plots the set of feasible outcomes $(\bar{p}, s)$ among all possible segmentations. An immediate consequence of Proposition 5 is a characterization of surplus division across possible segmentations. Figure IVB plots this set, where the horizontal axis is the surplus on the demand side and the vertical axis represents total transaction revenue.

From Proposition 5, if the platform's objective is to maximize total sales revenue, the optimal market segmentation leaves riders zero surplus and generates revenue $\int_0^\tau F_0^{-1}(1-x)\,\mathrm{d}x$, which is exactly the same as that under first-degree price discrimination (point A in Figure IVB). Therefore, when the platform only cares about sales revenue, the thickness constraint is, in fact, irrelevant, and the platform can extract all the surplus.

Alternatively, the platform might also have concern over rider surplus, perhaps to attract users and expand its network. In this case, the platform's objective function would be increasing in both sales revenue and rider surplus, which implies that the optimal market segmentation must generate a surplus division on the line segment connecting A and B in Figure IVB.

6 Application III: Econometrics

In this section, we apply our main result to subjects in econometrics. To this end, consider a random vector $(Y, X) \in \mathbb{R} \times \mathbb{R}^K$ on an underlying probability space with probability measure

$\mathbb{P}$.[20] Our main interest is about the conditional distribution of Y given realizations $X = x \in \mathbb{R}^K$, which we denote by $F_{Y|X=x}$. Note that for each $x \in \mathbb{R}^K$, $F_{Y|X=x}$ can be regarded as a realized posterior induced by the joint distribution of (Y, X).

6.1 Model Misspecification and Partial Identification in Quantile Regression

Koenker and Bassett Jr (1978) introduced the econometric approach of quantile regression, which models the quantiles of the conditional distribution of a response variable as a function of observed covariates. Our characterization can evaluate whether a presumed model for the conditional quantiles is mis-specified.

In particular, consider a response variable Y_i and K-dimensional covariate vector X_i. The observations $\{Y_i, X_i\}_{i=1}^N$ are independently and identically drawn. The marginal distribution of Y_i is either known from the literature or can be correctly estimated to be $F_0 \in \mathcal{F}$. Consider now the τ-quantile function $g_\tau : \mathbb{R}^K \to \mathbb{R}$ such that

$$g_\tau(x) \in \mathbb{Q}^\tau(F_{Y_i|X_i=x}),\ \forall x \in \mathbb{R}^K$$

Quantile regression aims to estimate the function g_τ using the sample. To facilitate the analysis and maintain tractability, econometricians often impose some further assumptions on the functional form of g_τ. A commonly used model is the linear model $g_\tau(x) = (1, x')\beta$, for some $\beta \in \mathbb{R}^{K+1}$. However, these models may potentially be mis-specified. Theorem 1 provides a simple test for model mis-specification.

Consider the (potentially mis-specified) linear quantile regression model $g_\tau(x) = (1, x')\beta$. The estimand under this model solves the minimization problem

$$\min_{\beta \in \mathbb{R}^{K+1}} \mathbb{E}[\rho_\tau(Y_i - (1, X_i')\beta)], \tag{12}$$

where $\rho_\tau(u) := u(\tau - \mathbf{1}\{u < 0\})$ is the tilted absolute value function.

From Theorem 1, the linear model is correctly specified under $\mathbb{P}$ only if the distribution of $(1, X_i')\beta^*$ is in $\mathcal{I}(\underline{F}_0^\tau, \overline{F}_0^\tau)$, where β^* is the solution to (12). An econometrician could test for model mis-specification using only knowledge of the marginal of Y_i. If the empirical distribution of $(1, X_i')\beta^*$ fell outside the interval $\mathcal{I}(\underline{F}_0^\tau, \overline{F}_0^\tau)$, there would be strong evidence of mis-specification. A comparison of the empirical quantiles or a Kolmogorov–Smirnov test are two ways to implement the evaluation. The reliance on information from just the marginal of Y_i allows one to bypass estimation of the joint distribution of (Y_i, X_i), which may

[20] Throughout this section, we hold fix this probability space and assume that it is rich enough relative to the random variable Y in the sense of definition 2 of Yang (2020). That is, the probability space restricted to any pre-image of Y is isomorphic to a unit interval with the Lebesgue measure.

be computationally demanding. Meanwhile, as the test does not exploit any conditioning information from X_i, which would have restricted the set of eligible models, the test has the advantage of presenting low Type-I error, but also the disadvantage of having low power.

When the number of covariates $K = 1$ and the marginals of both X_i and Y_i are known, an econometrician can go beyond evaluating model mis-specification to partially identifying the quantile function $g_\tau(X_i)$. Suppose now that $X_i \sim G$ and $Y_i \sim F_0$. Taking a concrete example, one might have Y_i representing income and X_i standing for years of schooling, but the two variables are potentially from two non-overlapping samples of the population. Estimation of economic models that involve Y_i and X_i originating from different samples is part of the econometrics of *data combination* (Ridder and Moffitt 2007).

If g_τ is known to be increasing (such as wages increasing in years of schooling), then, for any $\omega \in \mathbb{R}$, the probability that the conditional quantile registers at or below ω, given by $\mathbb{P}(g_\tau(X_i) \leq \omega)$, is simply $G(g_\tau^{-1}(\omega))$. As a result, by Theorem 1, it must be that $G \circ g_\tau \in \mathcal{I}(\underline{F}_0^\tau, \overline{F}_0^\tau)$, and hence, for all $\omega \in \mathbb{R}$,

$$(\underline{F}_0^\tau)^{-1} \circ G(\omega) \leq g_\tau(\omega) \leq (\overline{F}_0^\tau)^{-1} \circ G(\omega), \tag{13}$$

for all $\omega \in \mathbb{R}$. Proposition 6 below formalizes this observation and provides a non-parametric partial identification of the function g_τ.

Proposition 6 (Identification Set)**.** *For any $\tau \in (0,1)$ and for any increasing function $g_\tau : \mathbb{R} \to \mathbb{R}$, the following are equivalent:*

1. *There exists a random variable X such that the marginal of X is G and $g_\tau(X) \in \mathbb{Q}^\tau(F_{Y_1|X})$ with probability 1.*
2. *$(\underline{F}_0^\tau)^{-1} \circ G(\omega) \leq g_\tau(\omega) \leq (\overline{F}_0^\tau)^{-1} \circ G(\omega)$, for all $\omega \in \mathbb{R}$.*

Proof. The proof for 1 implying 2 follows immediately from Theorem 1 and the fact that g_τ is increasing. To see that 2 implies 1, consider any increasing function g_τ satisfying (13). Let $H(\omega) := G(g_\tau^{-1}(\omega))$ for all $\omega \in \mathbb{R}$. Then, by Theorem 1, $H \in \mathcal{I}(\underline{F}_0^\tau, \overline{F}_0^\tau)$ implies that there exists a signal $\mu \in \mathcal{M}$ and a selection rule $r \in \mathcal{R}$ such that $H(\omega) = H^\tau(\omega|\mu, r)$ for all $\omega \in \mathbb{R}$. As $\omega \in \mathbb{R}$, theorem 2 of Yang (2020) ensures that there exists a random variable $\widetilde{X}$ such that $\mathbb{P}(F_{Y_1|\widetilde{X}} \in A) = \mu(A)$ for all measurable $A \subseteq \mathcal{F}$, and

$$\mathbb{P}(\tilde{g}_\tau(\widetilde{X}) \leq \omega) = H(\omega),$$

for all $\omega \in \mathbb{R}$, with $\tilde{g}_\tau$ being an increasing function such that $\tilde{g}_\tau(x) \in \mathbb{Q}^\tau(F_{Y_1|\widetilde{X}=x})$ for all $x \in \mathbb{R}$. Now define a random variable $X := g_\tau^{-1}(\tilde{g}_\tau(\widetilde{X}))$. We claim that the marginal of X

is G and that $g_\tau(X) \in \mathbb{Q}^\tau(F_{Y_1|X})$ with $\mathbb{P}$-probability 1. Indeed, since both $\tilde{g}_\tau$ and g_τ are increasing,

$$\mathbb{P}(X \leq x) = \mathbb{P}(g_\tau^{-1}(\tilde{g}_\tau(\widetilde{X})) \leq x) = \mathbb{P}(\tilde{g}_\tau(\widetilde{X}) \leq g_\tau(x)) = H(g_\tau(x)) = G(x).$$

In the meantime, since $g_\tau^{-1} \circ \tilde{g}_\tau$ is increasing, it must be that $\mathbb{Q}^\tau(F_{Y_1|\widetilde{X}=x}) = \mathbb{Q}^\tau(F_{Y_1|X=g_\tau^{-1}\circ\tilde{g}_\tau(x)})$ for $\mathbb{P} \circ \widetilde{X}^{-1}$-almost all $x \in \mathbb{R}$. Together with $g_\tau(X) = \tilde{g}_\tau(\widetilde{X})$, it then follows that $g_\tau(X) \in \mathbb{Q}^\tau(F_{Y_1|X})$ with $\mathbb{P}$-probability 1. This completes the proof.

■

Proposition 5 provides a complete characterization of the identification set of g_τ. That is, under the given assumptions (i.e., only the marginals of Y_i, X_i are known and g_τ is only known to be increasing), the quantile function g_τ must satisfy (13). Conversely, for any function g_τ satysfying (13), there exists a model that meets the given assumptions and induces a quantile function g_τ. This identification result requires neither parametric assumptions, nor knowledge about the joint distribution, except for monotonicity of the quantile function. It allows an econometrician to make inferences on the conditional distribution of, say, income on schooling, just from knowing the marginal distributions of wages and school years, with the two potentially being measured from different population samples.

6.2 Inferences of Joint Distributions from Marginals

As hinted above, one common obstacle faced by econometricians is the lack of information about the joint distribution, even though information about the marginals is available. Specifically, given two random variables Y, X, with known marginals F_0 and G, respectively, what can one infer about their joint?

Horowitz and Manski (1995) provide a characterization when F_0 has a positive density on its support and when X is binary, which might refer to the realization of an event that contaminates the dataset, or enrollment in the prescribed treatment of an experiment. If $X \in \{0, 1\}$ and $\mathbb{P}(X = 1) = p \in (0, 1/2]$, the authors provide sharp bounds on the conditional distributions. In particular, for any $\tau \in (0, 1)$, they show that $\mathbb{Q}^\tau(F_{Y|X=1}) \subseteq [F_0^{-1}(\tau p), F_0^{-1}(\tau p + 1 - p)]$. Moreover, for each of the two bounds $F_0^{-1}(\tau p)$ and $F_0^{-1}(\tau p + 1 - p)$, there exists a joint distribution that attains the bound. Our Theorem 1 extends this result by demonstrating that *any* distribution within these bounds is attainable by some joint distribution of Y and X, with X being binary, as stated below.

Proposition 7. *For any $\tau \in (0, 1)$, for any random variable Y with distribution F_0, and for any $\omega \in [F_0^{-1}(\tau p), F_0^{-1}(\tau p + 1 - p)]$, there exists a random variable $X \in \{0, 1\}$ with*

$\mathbb{P}(X=1)=p$ *such that* $\omega \in \mathbb{Q}^\tau(F_{Y|X=1})$.

Proof. If $\omega \leq (\overline{F}_0^\tau)^{-1}(p)$, let

$$H_p(y) := \begin{cases} 0, & \text{if } y < \omega \\ p, & \text{if } y \in [\omega, (\overline{F}_0^\tau)^{-1}(p)) \\ 1, & \text{if } y \geq (\overline{F}_0^\tau)^{-1}(p) \end{cases},$$

for all $y \in \mathbb{R}$. Meanwhile, if $\omega > (\overline{F}_0^\tau)^{-1}(p)$, let

$$H_{1-p}(y) := \begin{cases} 0, & \text{if } y < (\underline{F}_0^\tau)^{-1}(1-p) \\ p, & \text{if } y \in [(\underline{F}_0^\tau)^{-1}(1-p), \omega) \\ 1, & \text{if } y \geq \omega \end{cases},$$

for all $y \in \mathbb{R}$. Note that $F_0^{-1}(\tau p) = (\underline{F}_0^\tau)^{-1}(p)$ and $F_0^{-1}(\tau p + 1 - p) = (\overline{F}_0^\tau)^{-1}(1-p)$. Therefore, if $\omega \leq (\overline{F}_0^\tau)^{-1}(p)$, then $H_p \in \mathcal{I}(\underline{F}_0^\tau, \overline{F}_0^\tau)$; while if $\omega > (\overline{F}_0^\tau)^{-1}(p)$, then $H_{1-p} \in \mathcal{I}(\underline{F}_0^\tau, \overline{F}_0^\tau)$. By Theorem 1, there always exists some $H \in \mathcal{H}_\tau$ with binary support that assigns probability p to ω, which, in turn, implies that there exists a random variable $X \in \{0,1\}$ with $\mathbb{P}(X=1)=p$ such that $\omega \in \mathbb{Q}^\tau(F_{Y|X=1})$, as desired. ■

In the meantime, Cross and Manski (2002) discuss identification of so-called *long* regressions when the *short* conditional distributions are known, but the *long* ones are not. The authors describe an environment in which each member of a population is associated with a tuple (Y, X, Z), such that $Y \in \mathbb{R}$, X takes values in a finite dimensional Euclidean space, and Z belongs to a K-element finite set. The issue at hand is identification of the long regression $\mathbb{E}[Y|X,Z]$ when the short conditional distributions $\mathbb{P}(Y|X)$ and $\mathbb{P}(Z|X)$ are known, but the long conditional distribution $\mathbb{P}(Y|X,Z)$ is unknown. The language "long" and "short" borrows from Goldberger (1991). In their article, the authors give bounds on $(\mathbb{E}[Y|X, Z=z_k])_{k=1}^K$, though the bounds might not be sharp. Only in the special case when both $K=2$ and the response variable $Y \in \{0,1\}$, do the authors completely identify the set of $(\mathbb{E}[Y|X, Z=z_k])_{k=1}^K$. Our Theorem 1 complements this result: The conditional quantiles, $(\mathbb{Q}^\tau(F_{Y|X,Z=z_k}))_{k=1}^K$, are completely identified for all $K < \infty$, even if $|\text{supp}(Y)| = \infty$.

Proposition 8. *Let* $P_0 := 0$ *and let* $P_k := \sum_{j=1}^k p_k$ *for all* $k \in \{1, \ldots, K\}$. *For any random variable* Y *with distribution* F_0, *for any* $\tau \in (0,1)$, *and for any vector* $\boldsymbol{q} = (q_k)_{k=1}^K \in \mathbb{R}^K$ *with* $q_1 \leq \ldots \leq q_K$, *the following are equivalent.*

1. *There exists a random variable* Z *with support* $\{z_k\}_{k=1}^K$ *and* $\mathbb{P}(Z=z_k)=p_k$ *such that* $q_k \in \mathbb{Q}^\tau(F_{Y|Z=z_k})$ *for all* $k \in \{1, \ldots, K\}$.

2. $q_k \in [(\underline{F}_0^\tau)^{-1}(P_k), (\overline{F}_0^\tau)^{-1}(P_{k-1}^+)]$.

Proof. To see that 1 implies 2, consider any such random variable Z. Let H be the CDF of the induced probability distribution over $\{q_k\}_{k=1}^K$. Theorem 1 implies that $H \in \mathcal{I}(\underline{F}_0^\tau, \overline{F}_0^\tau)$, which in turn implies 2.

Conversely, to see that 2 implies 1, consider any $\boldsymbol{q} = (q_k)_{k=1}^K$ with $q_1 \leq \ldots \leq q_K$. Define a CDF H as follows:

$$H(\omega) := \begin{cases} 0, & \text{if } \omega < q_1 \\ P_k, & \text{if } \omega \in [q_{k-1}, q_k),\ k \in \{2, \ldots, K\} \\ 1, & \text{if } \omega \geq q_K \end{cases},$$

for all $\omega \in \mathbb{R}$. Then, 2 implies $H \in \mathcal{I}(\underline{F}_0^\tau, \overline{F}_0^\tau)$. By Theorem 1, $H \in \mathcal{H}_\tau$, and thus there exists a random variable Z with $\text{supp}(Z) = \{z_k\}_{k=1}^K$ and $\mathbb{P}(Z = z_k) = p_k$ such that $q_k \in \mathbb{Q}^\tau(F_{Y|Z=z_k})$ for all $k \in \{1, \ldots, K\}$, as desired. ■

It is noteworthy that although the statement of Proposition 8 does not include the control variable X for the short regression, there are no restrictions on the marginal of Y. Therefore, the conditional quantiles $(\mathbb{Q}^\tau(F_{Y|X,Z=z_k}))_{k=1}^K$ can be identified by applying Proposition 8 to the conditional distribution of Y given each realization of X. Moreover, the monotonicty restriction in Proposition 8 is merely a normalization. The identification set of $(\mathbb{Q}^\tau(F_{Y|X,Z=z_k}))_{k=1}^K$ can be obtain by applying permutations on the result of Proposition 8, as illustrated by Figure V for the case of $K = 2$, $F_0 = U$, and $\tau = 1/2$.

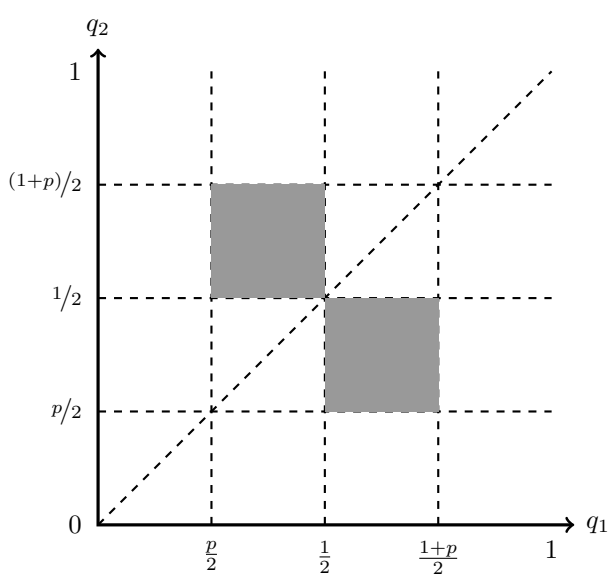


Figure V
Identification Set of $(\mathbb{Q}^{1/2}(F_{Y|Z=z_k}))_{k=1}^2$

7 Application IV: Finance and Accounting

7.1 Macroprudential Policy

Macroprudential policy deals with regulatory practices that aim at ensuring, as best as possible, the stability of the financial system as a whole (Galati and Moessner 2013). A commonly suggested policy suitable to this aim is to link a large financial institution's required amount of equity capital to its contribution to systemic risk, which is the risk of the entire financial system collapsing (Brunnermeier, Gorton, and Krishnamurthy 2012). Several measures of systemic risk abound, but a popular one is the Conditional Value at Risk from Adrian and Brunnermeier (2016), denoted $CoVaR$. The authors define this measure as the Value at Risk (VaR) of the financial system, conditional on a particular institution being under financial distress. In this context, the VaR would be the loss in the market value of the financial system that is exceeded with a certain (tail) probability (Duffie and Pan 1997).

But $CoVaR$ is inherently a quantile of a conditional probability distribution. Our characterization lets us describe the set of all possible $CoVaR$—and, hence, equity capital requirements—of a financial institution without knowing the exact correlation structure between the institution's returns and the system's returns. Knowing this range informs the regulator whether the macroprudential policy is reasonable or even feasible given the perceived appetite of the equity market to supply capital.

Specifically, let $R \sim F_0$ be the return of the financial system. The VaR of the system is defined as $F_0^{-1}(\tau)$ for some $\tau \in (0,1)$ (e.g., $\tau = 0.05$). Furthermore, let $R_i \sim F_i$ be financial institution i's return and let $X_i := \mathbf{1}\{R_i \leq \underline{r}_i\}$ signify the event that institution i's return is below some threshold $\underline{r}_i < 0$, which puts the institution in financial distress. The $CoVaR$ of financial institution i would simply be a τ-quantile of the distribution of R conditional on $X_i = 1$. Suppose that a financial regulator requires each institution i to have equity capital $g_i(\omega)$ if its $CoVaR$ is ω, where g_i is a decreasing function. (See Orlov, Zryumov, and Skrzypacz (2020) for a microfoundation for this kind of capital requirement.) An immediate implication of Theorem 1 (more precisely, Proposition 7) is a complete characterization of all possible equity capital requirements of the financial institution that can arise under some correlation structure between R_i and R, given the marginals F_0 and F_i. According to Proposition 7, institution i is required to issue equity capital e under some correlation structure if and only if $e \in [g_i(F_0^{-1}(\tau p_i)), g_i(F_0^{-1}(\tau p_i + 1 - p_i))]$, where p_i denotes the probability of $X_i = 1$ under F_i.

7.2 Classification Shifting

McVay (2006) describes a management tool to manipulate accounting earnings that involves deliberately misclassifying items within a firm's income statement. McVay refers to this practice as *classification shifting*, and she finds evidence consistent with managers moving expenses from the category of core expenses (e.g., cost of goods sold) to the category of special items (e.g., fines). Mangers are thought to engage in this conduct to overstate "core" earnings and meet Wall Street analyst earnings forecasts, as special items tend to be excluded from analysts' definitions of earnings. Fan, Barua, Cready, and Thomas (2010) document further evidence of classification shifting, and Dye (2002) presents a theoretical model of classification manipulation.

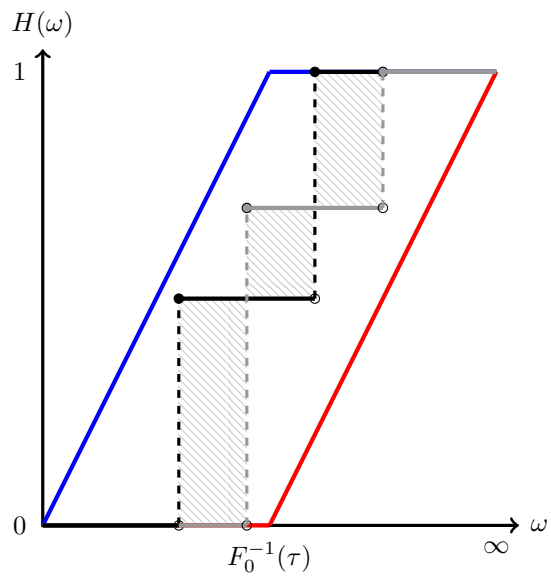


Figure VI
A Measure of Misclassification

With our result, we can provide a necessary condition for an auditor to undertake an inspection for classification shifting. Consider the following scenario: A manager classifies an extensive set of expenses into categories (e.g., core versus special items) and is requested to report a τ-quantile of each category to an auditor. Each expense belongs to a rightful category consistent with Generally Accepted Accounting Principals (GAAP). A certain dollar threshold of misclassification is considered material and constitutes accounting fraud. The auditor can observe the distribution of all expenses (in dollars), but a costly audit is required to verify each expense's classification. The problem for the auditor is to determine whether a closer inspection of the manager's classification is warranted.

Specifically, suppose the dollar threshold of misclassification to reach accounting fraud is

$K > 0$. Denote the distribution of expenses by F_0. Suppose that the auditor can select any $\tau \in (0,1)$ and request the manager to report a τ-quantile of each category. Then, for any $\tau \in (0,1)$, a report under any classification of spending induces a distribution of posterior τ-quantiles.

Let G_τ denote the distribution of τ-quantiles under the correct classification. If H is the distribution of τ-quantiles induced by the manager's classification, then the amount of misclassification, in the unit of dollars, can be measured as

$$\int_0^1 \left| G_\tau^{-1}(q) - H^{-1}(q) \right| \mathrm{d}q.$$

Figure VI illustrates a distribution of quantiles induced by a classification with two categories. The manager's classification is represented in black, and the correct classification is in gray.

By Theorem 1, both the distribution of τ-quantiles induced by the manager's classification and by the correct classification must reside within the first-order stochastic dominance interval $\mathcal{I}(\underline{F}_0^\tau, \overline{F}_0^\tau)$. As a result, whenever

$$\sup_{\tau \in [0,1]} \int_0^\infty \left| \overline{F}_0^\tau(\omega) - \underline{F}_0^\tau(\omega) \right| \mathrm{d}\omega \leq K, \tag{14}$$

no audit is warranted, since the left-hand side of (14) is the largest possible amount of misclassification.

8 Conclusion

We characterize the distributions of all possible posterior quantiles in a general environment. Unlike the distributions of posterior means, which are known to be mean-preserving contractions of the prior, the distributions of posterior quantiles reside between two first-order stochastic dominance bounds that are truncations of the prior. We apply this characterization to many economic scenarios, ranging across political economy, Bayesian persuasion, industrial organization, econometrics, finance, and accounting.

Other applications involving posterior quantiles undoubtedly exist. When consumers' values or firms' marginal costs follow distributions, different points on the inverse supply and demand curves are quantiles, which opens the door to further applications in consumer or firm theory. Inequality is often measured as an upper percentile of the wealth or income distribution, making it eligible for analysis. Likewise, settings in which threshold behavior is important, such as in theories of bank runs, protests, fads and fashions, or tipping points, are yet other directions for future work.

Appendix

A.1 Proof of Lemma 1

Since the CDF of the uniform distribution on $[0,1]$ is in $\mathcal{F}$, $\mathcal{H}_\tau = \mathcal{I}(\underline{F}_0^\tau, \overline{F}_0^\tau)$ for all $F_0 \in \mathcal{F}$ implies $\mathcal{H}_\tau^* \supseteq \mathcal{I}_\tau^*$.

Conversely, suppose that $\mathcal{H}_\tau^* \supseteq \mathcal{I}_\tau^*$. Consider any $H \in \mathcal{I}(\underline{F}_0^\tau, \overline{F}_0^\tau)$. Let $\widetilde{H}(q) := H(F_0^{-1}(q))$ for all $q \in \mathbb{R}$. Then $\widetilde{H} \in \mathcal{I}_\tau^*$. Therefore, there exists $\tilde{\mu} \in \mathcal{M}(U)$ and $\tilde{r} \in \mathcal{R}$ such that

$$\widetilde{H}(q) = H^\tau(q|\tilde{\mu}, \tilde{r}) = \int_{\mathcal{F}} \tilde{r}((-\infty, q]|F)\tilde{\mu}(\mathrm{d}F),$$

for all $q \in [0,1]$. For any $F_0 \in \mathcal{F}$, define μ and r as

$$\mu(A) := \tilde{\mu}(\{F \in \mathcal{F} | F \circ F_0 \in A\}),$$

for all measurable $A \subseteq \mathcal{F}$, and

$$r((-\infty, \omega]|F, \tau) := \tilde{r}((-\infty, F_0(\omega)]|F \circ F_0^{-1}, \tau),$$

for all $\omega \in \mathbb{R}$, for all $\tau \in (0,1)$, and for all $F \in \mathcal{F}$. We claim that $\mu \in \mathcal{M}(F_0)$ and $r \in \mathcal{R}$. Indeed, for any measurable $A \subseteq \mathcal{F}$, $\mu(A) = \tilde{\mu}(\{F \in \mathcal{F} | F \circ F_0 \in A\}) \geq 0$. Meanwhile, $\mu(\mathcal{F}) = \tilde{\mu}(\{F \in \mathcal{F} | F \circ F_0 \in \mathcal{F}\}) = \tilde{\mu}(\mathcal{F}) = 1$. Furthermore, for any measurable set $A \subseteq \mathcal{F}$, let

$$F_0^{-1} \circ A := \{F_0^{-1} \circ F | F \in A\},$$

and note that $F \circ F_0 \in A$ if and only if $F \in F_0^{-1} \circ A$ for all $F \in \mathcal{F}$. Thus, for any disjoint collection of measurable sets $\{A_n\} \subseteq \mathcal{F}$,

$$\begin{aligned}
\mu\left(\bigcup_{n=1}^\infty A_n\right) = \tilde{\mu}\left(\left\{F \in \mathcal{F} \middle| F \circ F_0 \in \bigcup_{n=1}^\infty A_n\right\}\right) &= \tilde{\mu}\left(\left\{F \in \mathcal{F} \middle| F \in F_0^{-1} \circ \bigcup_{n=1}^\infty A_n\right\}\right) \\
&= \sum_{n=1}^\infty \tilde{\mu}(F_0^{-1} \circ A_n) \\
&= \sum_{n=1}^\infty \tilde{\mu}(\{F \in \mathcal{F} | F \circ F_0 \in A_n\}) \\
&= \sum_{n=1}^\infty \mu(A_n).
\end{aligned}$$

Consequently, μ is indeed a probability measure on $\mathcal{F}$. In the meantime, for any $F \in \mathcal{F}$,

$$r((-\infty, F^{-1}(\tau))|F, \tau) = \tilde{r}((-\infty, F_0(F^{-1}(\tau))|F \circ F_0^{-1}, \tau) = 0$$

and

$$r((-\infty, F^{-1}(\tau^+)]|F, \tau) = \tilde{r}((-\infty, F_0(F^{-1}(\tau^+))|F \circ F_0^{-1}, \tau) = 1.$$

Thus, $\text{supp}(r(\cdot|F,\tau)) = \mathbb{Q}^\tau(F)$ for all $F \in \mathcal{F}$ and for all $\tau \in (0,1)$, and hence $r \in \mathcal{R}$.

In addition, for any $\omega \in \mathbb{R}$,

$$\int_{\mathcal{F}} F(\omega)\mu(\mathrm{d}F) = \int_{\mathcal{F}} F(F_0(\omega))\tilde{\mu}(\mathrm{d}F) = F_0(\omega),$$

which in turn implies that $\mu \in \mathcal{M}(F_0)$.

As a result, for any $\omega \in \mathbb{R}$,

$$\begin{aligned} H(\omega) = \widetilde{H}(F_0(\omega)) = \int_{\mathcal{F}} \tilde{r}((-\infty, F_0(\omega)]|F,\tau)\tilde{\mu}(\mathrm{d}F) &= \int_{\mathcal{F}} \tilde{r}((-\infty, F_0(\omega)]|F \circ F_0^{-1},\tau)\mu(\mathrm{d}F) \\ &= \int_{\mathcal{F}} r((-\infty,\omega]|F,\tau)\mu(\mathrm{d}F) \\ &= H^\tau(\omega|\mu, r). \end{aligned}$$

Therefore, $H \in \mathcal{H}_\tau$. This completes the proof. ■

A.2 Proof of Lemma 2

Embed $\mathcal{I}_\tau^* \subseteq \mathcal{F}$ into the collection $L^1([0,1])$ of integrable functions on $[0,1]$. Note that $\mathcal{I}_\tau^*$ is a convex subset of a normed linear space $L^1([0,1])$. Consider any $H \in \mathcal{I}_\tau^*$ that takes the form of (3), and any $\widehat{H} \in L^1([0,1])$ such that $\widehat{H}(\tilde{\omega}) \neq 0$ for some $\tilde{\omega} \in [0,1]$. Suppose that $H(\tilde{\omega}) \in \{\underline{U}^\tau(\tilde{\omega}), \overline{U}^\tau(\tilde{\omega})\}$. Then clearly either $H(\tilde{\omega}) + \widehat{H}(\tilde{\omega}) > \underline{U}^\tau(\tilde{\omega})$ or $H(\tilde{\omega}) - \widehat{H}(\tilde{\omega}) < \overline{U}^\tau(\tilde{\omega})$ and hence, either $H + \widehat{H} \notin \mathcal{I}_\tau^*$ or $H - \widehat{H} \notin \mathcal{I}_\tau^*$. Meanwhile, suppose that $\tilde{\omega} \in [\underline{x}_i, \overline{x}_i)$ for some $i \in I$ or $\tilde{\omega} \in [\underline{y}_j, \overline{y}_j)$ for some $j \in J$. If either $H + \widehat{H} \notin \mathcal{F}$ or $H - \widehat{H} \notin \mathcal{F}$, then clearly either $H + \widehat{H} \notin \mathcal{I}_\tau^*$ or $H - \widehat{H} \notin \mathcal{I}_\tau^*$. If, on the other hand, both $H + \widehat{H}$ and $H - \widehat{H}$ are in $\mathcal{F}$, then it must be that either $H(\omega) + \widehat{H}(\omega) = \underline{U}^\tau(\underline{x}_i) + \widehat{H}(\tilde{\omega}) > \underline{U}^\tau(\underline{x}_i)$ for all $\omega \in [\underline{x}_i, \overline{x}_i)$, or $H(\omega) - \widehat{H}(\omega) = \overline{U}^\tau(\overline{y}_j) - \widehat{H}(\tilde{\omega}) < \overline{U}^\tau(\overline{y}_j)$, for all $\omega \in [\underline{y}_j, \overline{y}_j)$. Therefore, there must exist $\hat{\omega} \in \mathbb{R}$ such that either $H(\hat{\omega}) + \widehat{H}(\hat{\omega}) \notin \mathcal{I}_\tau^*$ or $H(\hat{\omega}) - \widehat{H}(\hat{\omega}) \notin \mathcal{I}_\tau^*$.

Conversely, suppose that $H \in \mathcal{I}_\tau^*$ does not take form of (3). Then there exists $\underline{\omega} < \overline{\omega}$ and $\underline{\eta} < \overline{\eta}$ such that $H(\underline{\omega}^-) \leq \underline{\eta} \leq H(\underline{\omega})$, $H(\overline{\omega}^-) \leq \overline{\eta} \leq H(\overline{\omega})$; that $\overline{U}^\tau(\overline{\omega}) \leq \underline{\eta} < \overline{\eta} \leq \underline{U}^\tau(\underline{\omega})$; and that $\underline{\eta} < H(\omega) < \overline{\eta}$ for some $\omega \in (\underline{\omega}, \overline{\omega})$. Then, since the set of extreme points of nondecreasing functions that map from $[\underline{\omega}, \overline{\omega}]$ to $[\underline{\eta}, \overline{\eta}]$ must only take values in $\{\underline{\eta}, \overline{\eta}\}$ (see, for instance, lemma 2.7 of Börgers 2015), there exists a non-zero, integrable function $\widetilde{H} : [\underline{\omega}, \overline{\omega}] \to [\underline{\eta}, \overline{\eta}]$ such that both $H + \widetilde{H}$ and $H - \widetilde{H}$ are nondecreasing, right-continuous functions from $[\underline{\omega}, \overline{\omega}]$ to $[\underline{\eta}, \overline{\eta}]$. As a result, for any $\omega \in [\underline{\omega}, \overline{\omega}]$, it must be that

$$\max\{H(\omega) + \widetilde{H}(\omega), H(\omega) - \widetilde{H}(\omega)\} \leq \overline{\eta} \leq \underline{U}^\tau(\underline{\omega}) \leq \underline{U}^\tau(\omega) \tag{A.15}$$

and that

$$\min\{H(\omega) + \widetilde{H}(\omega), H(\omega) - \widetilde{H}(\omega)\} \geq \underline{\eta} \geq \overline{U}^\tau(\overline{\omega}) \geq \overline{U}^\tau(\omega), \tag{A.16}$$

for all $\omega \in [\underline{\omega}, \overline{\omega}]$. Now let $\widehat{H} : [0,1] \to \mathbb{R}$ be defined as

$$\widehat{H}(\omega) := \begin{cases} \widetilde{H}(\omega), & \text{if } \omega \in [\underline{\omega}, \overline{\omega}] \\ 0, & \text{otherwise} \end{cases},$$

for all $\omega \in [0,1]$. Clearly, $\widehat{H} \in L^1([0,1])$. Moreover, for any $\omega \in [0,1]$, from (A.15) and (A.16), together with $H \in \mathcal{I}_\tau^*$, it follows that

$$\overline{U}^\tau(\omega) \leq \min\{H(\omega) + \widehat{H}(\omega), H(\omega) - \widehat{H}(\omega)\} \leq \max\{H(\omega) + \widehat{H}(\omega), H(\omega) - \widehat{H}(\omega)\} \leq \underline{U}^\tau(\omega),$$

for all $\omega \in [0,1]$. Meanwhile, since $\underline{\eta} \in [H(\underline{\omega}^-), H(\underline{\omega})]$ and $\overline{\eta} \in [H(\overline{\omega}^-), H(\overline{\omega})]$, it must be that

$$H(\omega) + \widehat{H}(\omega) = H(\omega) - \widehat{H}(\omega) = H(\omega) \leq H(\underline{\omega}^-) \leq \underline{\eta},$$

for all $\omega \leq \underline{\omega}$; while

$$H(\omega) + \widehat{H}(\omega) = H(\omega) - \widehat{H}(\omega) = H(\omega) \geq H(\overline{\omega}) \geq \overline{\eta},$$

for all $\omega \geq \overline{\omega}$. As a result, both $H + \widehat{H}$ and $H - \widehat{H}$ are nondecreasing and right-continuous. It then follows that $H + \widehat{H} \in \mathcal{I}_\tau^*$ and $H - \widehat{H} \in \mathcal{I}_\tau^*$, and hence H is not an extreme point of $\mathcal{I}_\tau^*$. This completes the proof. ■

A.3 Proof of Theorem 1

To show that $\mathcal{H}_\tau \subseteq \mathcal{I}(\underline{F}_0^\tau, \overline{F}_0^\tau)$, consider any $H \in \mathcal{H}_\tau$. Let $\mu \in \mathcal{M}$ and any $r \in \mathcal{R}$ be a signal and a selection rule, respectively, such that $H^\tau(\cdot|\mu, r) = H$. By the definition of $H^\tau(\cdot|\mu, r)$, it must be that, for all $\omega \in \mathbb{R}$,

$$H(\omega|\mu, r) \leq \mu(\{F \in \mathcal{F} | F^{-1}(\tau) \leq \omega\}) = \mu(\{F \in \mathcal{F} | F(x) \geq \tau\}).$$

Now consider any $\omega \in \mathbb{R}$. Clearly, $\mu(\{F \in \mathcal{F} | F(\omega) \geq \tau\}) \leq 1$, since μ is a probability measure. Moreover, let $M_\omega^+(q) := \mu(\{F \in \mathcal{F} | F(\omega) \geq q\})$ for all $q \in [0,1]$. From (1), it follows that the left-limit of $1 - M_x^+$ is a CDF and a mean-preserving spread of a Dirac measure at $F_0(\omega)$. Therefore, whenever $\tau \geq F_0(\omega)$, then $M_\omega^+(\tau)$ can be at most $F_0(\omega)/\tau$ to have a mean of $F_0(\omega)$.[21] Together, this implies that $\mu(\{F \in \mathcal{F} | F(x) \geq \tau\}) \leq \underline{F}_0^\tau(\omega)$ for all $\omega \in \mathbb{R}$.

At the same time, by the definition of $H^\tau(\cdot|\mu, r)$, it must be that, for all $\omega \in \mathbb{R}$,

$$H^\tau(\omega^-|\mu, r) \geq \mu(\{F \in \mathcal{F} | F^{-1}(\tau^+) < \omega\}) = \mu(\{F \in \mathcal{F} | F(x) > \tau\}).$$

Now consider any $\omega \in \mathbb{R}$. Since μ is a probability measure, it must be that $\mu(\{F \in \mathcal{F} | F(\omega) > \tau\}) \geq 0$. Furthermore, let $M_\omega^-(q) := \mu(\{F \in \mathcal{F} | F(\omega) > q\})$ for all $q \in [0,1]$. From (1), it follows that $1 - M_x^-$ is a CDF and a mean-preserving spread of a Dirac measure at $F_0(\omega)$. Therefore, whenever $\tau \leq F_0(\omega)$, then $M_\omega^-(\tau)$ must be at least $(F_0(\omega) - \tau)/(1 - \tau)$ to have a mean of $F_0(\omega)$.[22] Together, this implies that $\mu(\{F \in \mathcal{F} | F(\omega) > \tau\}) \geq \overline{F}_0^\tau$ for all $\omega \in \mathbb{R}$, which, in turn, implies that $\overline{F}_0^\tau(\omega) \leq H^\tau(\omega^-|\mu, r) \leq H^\tau(\omega|\mu, r) \leq \underline{F}_0^\tau(\omega)$ for all $\omega \in \mathbb{R}$, as desired.

To prove that $\mathcal{I}(\underline{F}_0^\tau, \overline{F}_0^\tau) \subseteq \mathcal{H}_\tau$, by Lemma 1, it suffices to show that $\mathcal{I}_\tau^* \subseteq \mathcal{H}_\tau^*$. To this end, we first show that for any extreme point H of $\mathcal{I}_\tau^*$, there exists a signal $\tilde{\mu} \in \mathcal{M}(U)$ and a selection rule $\tilde{r} \in \mathcal{R}$ such that $H(\omega) = H^\tau(\omega|\mu, r)$ for all $\omega \in \mathbb{R}$. Consider any extreme point H of $\mathcal{I}_\tau^*$. By Lemma 2, H must take

[21] More specifically, to maximize the probability at τ, a mean-preserving spread of $F_0(\omega)$ must assign probability $F_0(\omega)/\tau$ at τ, and probability $1 - F_0(x)/\tau$ at 0.

[22] More specifically, to minimize the probability at τ, a mean-preserving spread of $F_0(x)$ must assign probability $(F_0(\omega) - \tau)/(1 - \tau)$ at 1, and probability $1 - (F_0(\omega) - \tau)/(1 - \tau)$ at 0.

the form of (3) for some $\underline{x}, \overline{x}, \underline{y}, \overline{y} \in [0,1]$ and countable sequences $\{\underline{x}_i, \overline{x}_i\}_{i\in I}$ and $\{\underline{y}_j, \overline{y}_j\}_{j\in J}$, such that $\underline{x} \leq \underline{x}_i \leq \overline{x}_i \leq \underline{x}_{i+1} \leq \overline{x} < \underline{y} \leq \underline{y}_j \leq \overline{y}_j \leq \underline{y}_{j+1} \leq \overline{y}$ for all $i \in I, j \in J$. Now define two classes of distributions, $\{\underline{U}^\omega\}_{\omega\in[0,\overline{x}]}$ and $\{\overline{U}^\omega\}_{\omega\in[\underline{y},1]}$, as follows:

$$\underline{U}^\omega(x) := \begin{cases} 0, & \text{if } x < \omega \\ \frac{\overline{x}}{\underline{y}-\tau+\overline{x}}, & \text{if } x \in [\omega, \tau) \\ \frac{x-\tau+\overline{x}}{1-\underline{y}+\overline{x}}, & \text{if } x \in [\tau, \underline{y}) \\ 1, & \text{if } x \geq \underline{y} \end{cases} ; \text{ and } \overline{U}^\omega(x) := \begin{cases} 0, & \text{if } x < \overline{x} \\ \frac{x-\overline{x}}{1-\underline{y}+\tau-\overline{x}}, & \text{if } x \in [\overline{x}, \tau) \\ \frac{\tau-\overline{x}}{1-\underline{y}+\tau-\overline{x}}, & \text{if } x \in [\overline{x}, \omega) \\ 1, & \text{if } x \geq \omega \end{cases}.$$

Since $\underline{U}^\tau(\overline{x}) = \overline{U}^\tau(\underline{y})$, it follows that $(1-\tau)\overline{x} = \tau(\underline{y}-\tau)$, and hence, $\underline{U}^\omega(x) = \tau$ for all $x \in [\omega, \underline{y}]$ and $\overline{U}^\omega(x) = \tau$ for all $x \in [\overline{x}, \omega)$. As a result, $\mathbb{Q}^\tau(\underline{U}^\omega) = [\omega, \underline{y}]$ for all $\omega \in [0, \overline{x}]$ and $\mathbb{Q}^\tau(\overline{U}^\omega) = [\overline{x}, \omega]$ for all $\omega \in [\underline{y}, 1]$. Moreover, for any $i \in I$ and for any $j \in J$, let $\underline{U}^i$ and $\overline{U}^j$ be defined as

$$\underline{U}^i(x) := \frac{1}{\overline{x}_i - \underline{x}_i} \int_{\underline{x}_i}^{\overline{x}_i} \underline{U}^\omega(x)\,\mathrm{d}\omega; \text{ and } \overline{U}^j(x) := \frac{1}{\overline{y}_j - \underline{y}_j} \int_{\underline{y}_j}^{\overline{y}_j} \overline{U}^\omega(x)\,\mathrm{d}\omega,$$

for all $x \in \mathbb{R}$. By construction, $\underline{U}^i, \overline{U}^j \in \mathcal{F}$ and $\overline{x}_i \in \mathbb{Q}^\tau(\underline{U}^i)$, $\underline{y}_j \in \mathbb{Q}^\tau(\overline{U}^j)$ for all $i \in I$ and $j \in J$. Next, for any $\omega \in \mathrm{supp}(H)$, let $F_\omega \in \mathcal{F}$ be defined as[23]

$$F_\omega(x) := \begin{cases} \underline{U}^\omega(x), & \text{if } \omega \in [0, \overline{x}] \setminus \cup_{i\in I} [\underline{x}_i, \overline{x}_i] \\ \underline{U}^i(x), & \text{if } \omega \in [\underline{x}_i, \overline{x}_i] \\ \overline{U}^\omega(x), & \text{if } \omega \in [\underline{y}, 1] \setminus \cup_{j\in J} [\underline{y}_j, \overline{y}_j] \\ \overline{U}^j(x), & \text{if } \omega \in [\underline{y}_j, \overline{y}_j] \end{cases},$$

for all $x \in \mathbb{R}$.

Now define $\tilde{\mu}$ as

$$\tilde{\mu}(\{F_\omega \in \mathcal{F} | \omega \leq x\}) := H(x),$$

for all $x \in \mathbb{R}$. By construction, $\mathrm{supp}(\tilde{\mu}) = \{\underline{U}^\omega\}_{\omega\in[0,\overline{x}]\setminus\cup_{i\in I}[\underline{x}_i,\overline{x}_i]} \cup \{\underline{U}^i\}_{i\in I} \cup \{\overline{U}^\omega\}_{\omega\in[\underline{y},1]\setminus\cup_{j\in J}[\underline{y}_i,\overline{y}_j]} \cup \{\overline{U}^j\}_{j\in J}$. Furthermore, for any $x \in [0,1]$,

$$\int_{\mathcal{F}} F(x)\tilde{\mu}(\mathrm{d}F) = \int_0^1 F_\omega(x) H(\mathrm{d}\omega) = x,$$

and hence $\tilde{\mu} \in \mathcal{M}(U)$. In the meantime, let $\tilde{r} : \mathcal{F} \to [0,1] \to \Delta(\mathbb{R})$ be defined as

$$\tilde{r}(F, \tau') := \begin{cases} \delta_{\{\max(\mathbb{Q}^{\tau'})\}}, & \text{if } F = F_\omega,\ \omega \in [\underline{y}, 1] \\ \delta_{\{\min(\mathbb{Q}^{\tau'})\}}, & \text{otherwise} \end{cases},$$

[23] As a convention, define $\underline{U}^i(x) := \underline{U}^\omega(x)$ for all x if $\underline{x}_i = \overline{x}_i = \omega$. Similarly, define $\overline{U}^j(x) := \overline{U}^\omega(x)$ for all x if $\underline{y}_j = \overline{y}_j = \omega$.

for all $F \in \mathcal{F}$ and for all $\tau' \in [0,1]$. Then, for all $x \in \mathbb{R}$,

$$H^{\tau}(x|\tilde{\mu},\tilde{r}) = \begin{cases} 0, & \text{if } x < 0 \\ \tilde{\mu}(F_\omega | F_\omega^{-1}(\tau) \leq x), & \text{if } x \in [0,\overline{x}) \\ \tilde{\mu}(F_\omega | F_\omega^{-1}(\tau^+) \leq x), & \text{if } x \in [\underline{y},1) \\ 1, & \text{if } x \geq 1 \end{cases}$$

for all $x \in \mathbb{R}$, and hence $H^{\tau}(\omega|\tilde{\mu},\tilde{r}) = H(\omega)$ for all $\omega \in \mathbb{R}$, as desired.

Lastly, let Γ be a collection of probability measures $\gamma \in \Delta(\mathbb{R} \times \mathcal{F})$ such that $\gamma(\{(\omega, F) \in \mathbb{R} \times \mathcal{F} | \omega \in \mathbb{Q}^{\tau}(F)\} = 1$ and

$$\int_{\mathbb{R}\times\mathcal{F}} F(\omega)\gamma(\mathrm{d}\tilde{\omega}, \mathrm{d}F) = U(\omega),$$

for all $\omega \in \mathbb{R}$. Define a linear functional $\Xi : \Gamma \to \mathcal{F}$ as

$$\Xi(\gamma)[\omega] := \gamma((-\infty,\omega],\mathcal{F}),$$

for all $\gamma \in \Gamma$ and for all $\omega \in \mathbb{R}$. Then, since for any $\widetilde{H}$ in the set of extreme points $\mathrm{ext}(\mathcal{I}_\tau^*)$ of $\mathcal{I}_\tau^*$, there exists $\tilde{\mu} \in \mathcal{M}(U)$ and $\tilde{r} \in \mathcal{R}$ such that $H^{\tau}(\omega|\tilde{\mu},\tilde{r}) = \widetilde{H}(\omega)$ for all $\omega \in \mathbb{R}$, it must be that $\mathrm{ext}(\mathcal{I}_\tau^*) \subseteq \Xi(\Gamma)$.

Now consider any $H \in \mathcal{I}_\tau^*$. Since $\mathcal{I}_\tau^*$ is a compact and convex set of a metrizable, locally convex topological space, Choquet's theorem implies that there exists a probability measure $\Lambda_H \in \Delta(\mathcal{I}_\tau^*)$ with $\Lambda_H(\mathrm{ext}(\mathcal{I}_\tau^*)) = 1$ such that

$$\int_{\mathcal{I}_\tau^*} \widetilde{H}(\omega)\Lambda_H(\mathrm{d}\widetilde{H}) = H(\omega),$$

for all $\omega \in \mathbb{R}$. Define a measure $\widetilde{\Lambda}_H$ by

$$\widetilde{\Lambda}_H(A) := \Lambda_H(\{\Xi(\gamma)|\gamma \in A\}),$$

for all measurable $A \subseteq \Gamma$. Since $\Lambda_H(\mathrm{ext}(\mathcal{I}_\tau^*)) = 1$ and $\mathrm{ext}(\mathcal{I}_\tau^*) \subseteq \Xi(\Gamma)$, $\widetilde{\Lambda}_H$ is a probability measure on Γ.

For any $\omega \in \mathbb{R}$ and for any measurable $A \subseteq \mathcal{F}$, let

$$\gamma((-\infty,\omega],A) := \int_\Gamma \tilde{\gamma}((-\infty,\omega],A)\widetilde{\Lambda}_H(\mathrm{d}\tilde{\gamma}),$$

and let $\mu(A) := \gamma(\mathbb{R}, A)$. By construction, for all $\omega \in \mathbb{R}$,

$$\int_{\mathcal{F}} F(\omega)\mu(\mathrm{d}F) = \int_\Gamma \left(\int_{\mathbb{R}\times\mathcal{F}} F(\omega)\tilde{\gamma}(\mathrm{d}\tilde{\omega}, \mathrm{d}F) \right) \widetilde{\Lambda}_H(\mathrm{d}\tilde{\gamma}) = U(\omega),$$

and hence $\mu \in \mathcal{M}(U)$. Furthermore, by the disintegration theorem (c.f., Çinlar 2010, theorem 2.18), there exists a transition probability $r : \mathcal{F} \to \Delta(\mathbb{R})$ such that $\gamma(\mathrm{d}\omega, \mathrm{d}F) = r(\mathrm{d}\omega|F)\mu(\mathrm{d}F)$. Since $\widetilde{\Lambda}_H(\Gamma) = 1$, it

must be that $r \in \mathcal{R}$. Finally, for any $\omega \in \mathbb{R}$, since Ξ is affine,

$$\begin{aligned} H^{\tau}(\omega|\mu, r) = \gamma((-\infty, \omega], \mathcal{F}) =& \Xi(\gamma)[\omega] \\ =& \int_{\Gamma} \Xi(\tilde{\gamma})[\omega] \widetilde{\Lambda}_H(\mathrm{d}\tilde{\gamma}) \\ =& \int_{\mathrm{ext}(\mathcal{I}_{\tau}^{*})} \widetilde{H}(\omega) \Lambda_H(\mathrm{d}\widetilde{H}) \\ =& H(\omega), \end{aligned}$$

as desired. This completes the proof. ■

A.4 Proof of Proposition 4

Fix any $\alpha \in [1/2, 1]$. We first prove that 1 implies 2. Consider any $\omega \in [\underline{\omega}(\alpha), \overline{\omega}(\alpha)]$. If $\omega \in \mathbb{Q}^{1/2}(F_0)$, then 2 must hold, since the map $\delta_{\{F_0\}} \in \mathcal{M}$ and the selection rule that selects ω with probability 1 induces a distribution of representatives that unanimously share an ideal position of ω. Now suppose that $\omega < F_0^{-1}(1/2)$. If the distribution of representatives' ideal positions is $\underline{F}_0^{1/2}$, then the share of representatives whose ideal positions are closer to ω than to $F_0^{-1}(1/2)$ would be $2F((F^{-1}(1/2) + \omega)/2)$, which, in turn, is at least α, as $\omega \geq \underline{\omega}(\alpha)$. Similarly, suppose that $\omega > F_0^{-1}(1/2^{+})$. If the distribution of representatives' ideal positions is $\overline{F}_0^{1/2}$, then the share of representatives whose ideal position is closer to ω than to $F^{-1}(1/2^{+})$ would be $2(1 - F((F^{-1}(1/2) + \omega)/2))$, which, in turn, is at least α, as $\omega \leq \overline{\omega}(\alpha)$. Therefore, by Theorem 1, 2 is satisfied for all $\omega \in [\underline{\omega}(\alpha), \overline{\omega}(\alpha)]$.

Conversely, to prove that 2 implies 1, fix any $\omega \in [0, 1]$ and suppose that there exists a map $\mu \in \mathcal{M}$ and a selection rule $r \in \mathcal{R}$ such that under $H^{1/2}(\cdot|\mu, r)$, the share of representatives with ideal positions closer to ω than to either of $F_0^{-1}(1/2)$ or $F_0^{-1}(1/2^{+})$ is at least α. That is, $H^{1/2}((F_0^{-1}(1/2) + \omega)/2|\mu, r) \geq \alpha$ if $\omega \leq F_0^{-1}(1/2)$ and $H^{1/2}((F_0^{-1}(1/2^{+}) + \omega)/2|\mu, r) \leq 1 - \alpha$ if $\omega \geq F_0^{-1}(1/2^{+})$. By Theorem 1, it then follows that

$$2F_0\left(\frac{F_0^{-1}(1/2) + \omega}{2}\right) \geq H^{1/2}\left(\frac{F_0^{-1}(1/2) + \omega}{2}\middle|\mu, r\right) \geq \alpha$$

if $\omega \leq F_0^{-1}(1/2)$, and

$$2F_0\left(\frac{F_0^{-1}(1/2) + \omega}{2}\right) - 1 \leq H^{1/2}\left(\frac{F_0^{-1}(1/2) + \omega}{2}\middle|\mu, r\right) \leq 1 - \alpha$$

if $\omega \geq F^{-1}(1/2^{+})$, which, in turn, implies $\underline{\omega}(\alpha) \leq \omega \leq \overline{\omega}(\alpha)$, as desired. ■

A.5 Proof of Proposition 5

Consider any market segmentation $\mu \in \mathcal{M}$ and let $(\bar{p}, s)$ be the induced average price and total surplus, respectively. Since the total surplus of this market is $\int_0^{\tau} F_0^{-1}(1 - x)\,\mathrm{d}x$, and since the smallest possible surplus is the one induced by random matching, (10) follows. Furthermore, since

$$s = \int_{\mathcal{F}} \left(\int_0^{\tau} F^{-1}(1 - x)\,\mathrm{d}x\right) \mu(\mathrm{d}F),$$

and since the function

$$\tau \mapsto \int_{\mathcal{F}} \left(\int_0^{\tau} F^{-1}(1-x)\,\mathrm{d}x \right) \mu(\mathrm{d}F)$$

is concave, it must be that

$$\frac{\int_0^1 F_0^{-1}(1-x)\,\mathrm{d}x - s}{1-\tau} \leq \bar{p} = \int_{\mathcal{F}} F^{-1}(1-\tau)\mu(\mathrm{d}F) \leq \frac{s}{\tau},$$

as desired.

Conversely, under any segmentation that induces the price distribution $\overline{F}_0^{(1-\tau)}$, the average price equals $\int_0^{\tau} F_0^{-1}(1-x)\,\mathrm{d}x/\tau$, and total surplus equals $\int_0^{\tau} F_0^{-1}(1-x)\,\mathrm{d}x$; whereas under any segmentation that induces the price distribution $\underline{F}_0^{(1-\tau)}$, the average price equals $(\int_0^1 F_0^{-1}(1-x)\,\mathrm{d}x - s)/(1-\tau)$, and total surplus equals $\int_0^{\tau} F_0^{-1}(1-x)\,\mathrm{d}x$.[24] Lastly, under random matching, the average price is $\int_0^1 F_0^{-1}(1-x)\,\mathrm{d}x$ and total surplus is $\tau \int_0^1 F_0^{-1}(1-x)\,\mathrm{d}x$. As a result, since the set given by (10) and (11) is convex, and since the extreme points of this set can be induced by some segmentation according to Theorem 1, every element of this set can be induced by some segmentation as well. ■

References

Adrian, T. and M. K. Brunnermeier (2016): "CoVaR," *American Economic Review*, 1705–1741.

Alexeev, B. and D. G. Mixon (2018): "An impossibility theorem for gerrymandering," *American Mathematical Monthly*, 125, 878–884.

Ali, S. N., N. Haghpanah, X. Lin, and R. Siegel (2022): "How to sell hard information," *Quarterly Journal of Economics*, 137, 619–678.

Anderson, S. P. and R. Renault (2006): "Advertising content," *American Economic Review*, 96, 93–113.

Aumann, R. J. and M. B. Maschler (1995): *Repeated games with incomplete information*, MIT Press.

Benoît, J.-P. and J. Dubra (2011): "Apparent overconfidence," *Econometrica*, 79, 1591–625.

Bergemann, D., B. Brooks, and S. Morris (2015): "The limits of price discrimination," *American Economic Review*, 105, 921–957.

Besley, T. and I. Preston (2007): "Electoral bias and policy choice: Theory and evidence," *Quarterly Journal of Economics*, 122, 1473–1510.

Black, D. (1958): *The Theory of Committees and Elections*, Springer.

[24] Recall that, for any $\omega \in \operatorname{supp}(\overline{F}_0^{(1-\tau)})$, ω is a $1-\tau$-quantile of the segment it belongs to. Hence, every rider in the same segment with value above ω must buy at that price. This, in turn, implies that the top $100 \cdot \tau$ percent of riders end up getting a ride.

Blackwell, D. (1953): "Equivalent comparisons of experiments," *Annals of Mathematical Statistics*, 24, 265–272.

Börgers, T. (2015): *An Introduction to the Theory of Mechanism Design*, Oxford University Press.

Bradbury, J. C. and W. M. Crain (2005): "Legislative district configurations and fiscal policy in American States," *Public Choice*, 125, 385–407.

Brunnermeier, M. K., G. Gorton, and A. Krishnamurthy (2012): "Risk topography," *NBER Macroeconomics Annual*, 26, 149–176.

Çinlar, E. (2010): *Probability and Stochastics*, Springer.

Chambers, C. P. (2007): "Ordinal aggregation and quantiles," *Journal of Economic Theory*, 137, 416–431.

Cho, S.-j. and J. Duggan (2009): "Bargaining foundations of the median voter theorem," *Journal of Economic Theory*, 144, 851–868.

Coate, S. and B. Knight (2007): "Socially optimal districting: A theoretical and empirical exploration," *Quarterly Journal of Economics*, 122, 1409–1471.

Condorelli, D. and B. Szentes (2022): "Buyer-optimal matching in two-sided platforms," Working Paper. University of Warwick, Coventry, UK.

Cross, P. J. and C. F. Manski (2002): "Regressions, short and long," *Econometrica*, 70, 357–368.

de Castro, L. and A. F. Galvao (2021): "Static and dynamic quantile preferences," *Economic Theory*, 1–33.

de Cornière, A. (2016): "Search advertising," *American Economic Journal: Microeconomics*, 8, 156–188.

Downs, A. (1957): *An Economic Theory of Democracy*, Harper & Row New York.

Duffie, D. and J. Pan (1997): "An overview of value at risk," *Journal of Derivatives*, 4, 7–49.

Dworczak, P. and G. Martini (2019): "The simple economics of optimal persuasion," *Journal of Political Economy*, 127, 1993–2048.

Dye, R. A. (2002): "Classifications manipulation and Nash accounting standards," *Journal of Accounting Research*, 40, 1125–1162.

Elliot, M., A. Galeotti, A. Koh, and W. Li (2022): "Market segmentation through information," Working Paper. London Business School, London, UK.

Ely, J. C. (2019): "A cake-cutting solution to gerrymandering," Working Paper. Northwestern University, Evanston, IL.

Fan, Y., A. Barua, W. M. Cready, and W. B. Thomas (2010): "Managing earnings using classification shifting: Evidence from quarterly special items," *Accounting Review*, 85, 1303–1323.

Friedman, J. N. and R. Holden (2020): "Optimal Gerrymandering in a competitive environment," *Economic Theory Bulletin*, 8, 347–367.

Friedman, J. N. and R. T. Holden (2008): "Optimal gerrymandering: Sometimes pack, but never crack," *American Economic Review*, 98, 113–44.

Fryer Jr, R. G. and R. Holden (2011): "Measuring the compactness of political districting plans," *Journal of Law and Economics*, 54, 493–535.

Galati, G. and R. Moessner (2013): "Macroprudential policy–A literature review," *Journal of Economic Surveys*, 27, 846–878.

Gentzkow, M. and E. Kamenica (2016): "A Rothschild-Stiglitz approach to Bayesian persuasion," *American Economic Review: Papers and Proceedings*.

Goldberger, A. S. (1991): *A Course in Econometrics*, Harvard University Press.

Gomberg, A., R. Pancs, and T. Sharma (2021): "Electoral maldistricting," Working Paper. ITAM, Mexico.

Guinsburg, P. V. and G. Q. Saraiva (2022): "Price steering in two-sided platforms," Working Paper. São Paulo University, São Paulo, Brazil.

Gul, F. and W. Pesendorfer (2010): "Strategic redistricting," *American Economic Review*, 100, 1616–41.

Haghpanah, N. and R. Siegel (2020): "Pareto improving segmentation of multi-product markets," Working Paper. Pennsylvania State University, State College, PA.

——— (forthcoming): "The limits of multi-product price discrimination," *American Economic Review: Insight*.

Hagiu, A. and B. Jullien (2011): "Why do intermediaries divert search," *RAND Journal of Economics*, 42, 337–362.

Hardy, G., J. E. Littlewood, and G. Pólya (1929): "Some simple inequalities satisfied by convex functions," *Messenger Math*, 58, 145–152.

Harsanyi, J. C. (1967-68): "Games of incomplete information played by Bayesian players, Part I, II, III," *Management Science*, 14, 159–182, 320–334, 486–502.

Horowitz, J. L. and C. F. Manski (1995): "Identification and robustness with contaminated and corrupted data," *Econometrica*, 281–302.

Ichihashi, S. (2020): "Online privacy and information disclosure by consumers," *American Economic Review*, 110, 569–595.

Kamenica, E. and M. Gentzkow (2011): "Bayesian persuasion," *American Economic Review*.

Kleiner, A., B. Moldovanu, and P. Strack (2021): "Extreme points and majorization: Economic applications," *Econometrica*, 89, 1557–1593.

Koenker, R. and G. Bassett Jr (1978): "Regression quantiles," *Econometrica*, 33–50.

Kolotilin, A. (2018): "Optimal information disclosure: A linear programming approach," *Theoretical Economics*, 13, 607–635.

Kolotilin, A., R. Corrao, and A. Wolitzky (2022): "Persuasion as matching," Working Paper. Massachusetts Institute of Technology, Cambridge, MA.

Kolotilin, A., M. Li, T. Mylovanov, and A. Zapechelnyuk (2017): "Persuasion of a privately informed receiver," *Econometrica*, 85, 1949–1964.

Kolotilin, A., T. Mylovanov, and A. Zapechelnyuk (forthcoming): "Censorship as optimal persuasion," *Theoretical Economics*.

Kolotilin, A. and A. Wolitzky (2020): "The economics of partisan gerrymandering," Working Paper. Massachusetts Institute of Technology, Cambridge, MA.

Krehbiel, K. (2010): *Pivotal Politics*, University of Chicago Press.

Lipnowski, E. and D. Ravid (2020): "Cheap talk with transparent motives," *Econometrica*, 88, 1631–60.

Manski, C. F. (1988): "Ordinal utility models of decision making under uncertainty," *Theory and Decision*, 25, 79–104.

McCarty, N., K. T. Poole, and H. Rosenthal (2001): "The hunt for party discipline in Congress," *American Political Science Review*, 95, 673–687.

——— (2009): "Does gerrymandering cause polarization?" *American Journal of Political Science*, 53, 666–680.

McGhee, E. (2014): "Measuring partisan bias in single-member district electoral systems," *Legislative Studies Quarterly*, 39, 55–85.

McVay, S. E. (2006): "Earnings management using classification shifting: An examination of core earnings and special items," *Accounting Review*, 81, 501–531.

Orlov, D., P. Zryumov, and A. Skrzypacz (2020): "Design of macro-prudential stress tests," Working paper. Stanford Univeristy, Stanford, CA.

Owen, G. and B. Grofman (1988): "Optimal partisan gerrymandering," *Political Geography Quarterly*, 7, 5–22.

Pegden, W., A. D. Procaccia, and D. Yu (2017): "A partisan districting protocol with provably nonpartisan outcomes," Working Paper. Carnegie Mellon University, Pittsburgh, PA.

Ridder, G. and R. Moffitt (2007): "The econometrics of data combination," *Handbook of Econometrics*, 6, 5469–5547.

Rostek, M. (2010): "Quantile maximization in decision theory," *Review of Economic Studies*, 77, 339–371.

Rothschild, M. and J. Stiglitz (1970): "Increasing risk: I. A Definition," *Journal of Economic Theory.*

Ryff, J. V. (1967): "Extreme points of some convex subsets of $L^1(0,1)$," *Proceedings of the American Mathematical Society*, 18, 1026–1034.

Shotts, K. W. (2001): "The effect of majority-minority mandates on partisan gerrymandering," *American Journal of Political Science*, 120–135.

Stephanopoulos, N. O. and E. M. McGhee (2015): "Partisan gerrymandering and the efficiency gap," *U. Chi. L. Rev.*, 82, 831.

Strassen, V. (1965): "The existence of probability measures with given marginals," *Annuals of Mathematical Statistics*, 36, 423–439.

Yang, K. H. (2020): "A note on generating arbitrary joint distributions using partitions," Working Paper. Yale University, New Haven, CT.

——— (2022): "Selling consumer data for profit: Optimal market-segmentation design and its consequences," *American Economic Review*, 112, 1364–1393.