Proceedings Article

Cost Function Market Makers for Measurable Spaces

YILING CHEN, Harvard University
MIKE RUBERRY, Harvard University
JENNIFER WORTMAN VAUGHAN, Microsoft Research, New York City & UCLA

We characterize cost function market makers designed to elicit traders' beliefs about the expectations of an infinite set of random variables or the full distribution of a continuous random variable. This characterization is derived from a duality perspective that associates the market maker's liabilities with market beliefs, generalizing the framework of Abernethy et al. [2011, 2013], but relies on a new subdifferential analysis. It differs from prior approaches in that it allows arbitrary market beliefs, not just those that admit density functions. This allows us to overcome the impossibility results of Gao and Chen [2010] and design the first automated market maker for betting on the realization of a continuous random variable taking values in $[0, 1]$ that has bounded loss without resorting to discretization. Additionally, we show that scoring rules are derived from the same duality and share a close connection with cost functions for eliciting beliefs.

Categories and Subject Descriptors: F.0 [**Theory of Computation**]: General; J.4 [**Social and Behavioral Sciences**]: Economics

General Terms: Theory, Economics

Additional Key Words and Phrases: Automated market makers, prediction markets

1. INTRODUCTION

There is a longstanding and widely held belief that markets aggregate information about uncertain future events [Ramsey 1926]. Colloquially, it would not be unusual to hear someone say "the market predicts that the price of corn will rise." In the past few decades, prediction markets have been designed specifically for the purpose of eliciting and aggregating information about events of interest from traders. Theoretical work [Ostrovsky 2012; Chen et al. 2012] has shown that such markets can produce an accurate consensus estimate, and empirical work [Berg et al. 2001; Wolfers and Zitzewitz 2004; Polgreen et al. 2007] has shown that markets can produce practically useful predictions in many settings.

Some prediction markets, like the (currently defunct) popular market Intrade, are implemented as continuous double auctions. This is fine when the number of traders is large and the space of available contracts small, but can lead to low liquidity when the number of traders is small or the space of contracts complex. For information elicitation, it is desirable to ensure that traders can always find a counter party with whom to trade to reveal their information. This motivates the recent line of research on the design of *automated market makers*, algorithmic agents who are always willing to trade at some price, adding liquidity to the market by taking on some risk [Hanson 2003]. Abernethy et al. [2011, 2013] characterized a broad family of *cost function market makers* that uniquely satisfy a set of

---

Authors' addresses: Y. C. and M. R., Harvard School of Engineering and Applied Sciences, 33 Oxford Street, Cambridge, MA; J. W. V., Microsoft Research, 102 Madison Ave., 12th Floor, New York, NY.
This research was partially supported by the NSF under grants IIS-1054911 and CCF-0953516. Any opinions, findings, conclusions, or recommendations expressed here are those of the authors alone.

desirable economic properties and can be used whenever predictions are being made over a finite outcome space (e.g., a race with n candidates and $n!$ possible orderings).

The framework of Abernethy et al. [2013] can be used to implement bets on the means of any finite number of continuous random variables, but for full distributions of continuous random variables, or even approximations of full distributions with restricted securities, not a single cost function has been demonstrated that has bounded worst-case loss without relying on some form of discretization. The most simple scenario to consider is a market taking bets on a continuous random variable taking values in $[0, 1]$. Traders are allowed to buy or sell securities over any subinterval $[a, b]$, with each contract worth \$1 if the value of the random variable is realized in the subinterval and \$0 otherwise. There have been several previous attempts to design cost function market makers for this scenario. Gao et al. [2009] observed that when the *logarithmic market scoring rule* (LMSR) [Hanson 2003, 2007] (a now-classic cost function market maker belonging to the family characterized by Abernethy et al. [2013]) is generalized to this setting, it suffers from unbounded loss. Gao and Chen [2010] later generalized this result to show that any reasonable market maker for continuous random variables on [0, 1] suffers unbounded worst-case loss under the implicit restriction that market beliefs can be represented by a probability density function. Othman and Sandholm [2010] successfully extended an existing constant-utility market maker to deal with countably infinite outcome spaces, but this market maker still has unbounded loss for continuous outcome spaces.

It is this stark contrast between success for finite outcome spaces and failure for continuous outcome spaces that motivates our development of a more general cost function market maker framework. Our framework enables the design of market makers for any measurable outcome space. With such a framework, we are able to develop cost function market makers that have bounded worst-case loss for betting on the interval $[0, 1]$ and other continuous spaces. One key difference between our framework and these prior efforts is that all prior work implicitly required that the market's beliefs must *always be representable by a probability density function*. In retrospect, there is no strong economic reason to require this; in fact, this intuitively should not be the case if traders purchase a large volume of securities that pay off only for a single outcome. Our cost function market makers allow the market beliefs to be represented as any probability measure. This relaxation makes it possible to achieve bounded loss over continuous outcome spaces, getting around the impossibility results of Gao and Chen [2010].

We characterize all cost function market makers for measurable outcome spaces that satisfy four desirable economic properties that are spiritually similar to those properties introduced by Abernethy et al. [2011, 2013]. Our characterization takes a similar duality perspective in describing a cost function's association of payoffs and beliefs. It can be shown that the family of market makers characterized by Abernethy et al. [2013] is a special case of ours, but our work is not a simple generalization. Our characterization relies on a new subdifferential analysis; one way to summarize our characterization is to say that the various relationships between beliefs and payoffs that incentivize traders to be accurate in a market are those encoded by strictly convex functions in Banach spaces and their subdifferentials.

Additionally, we establish a connection between our new characterization of cost functions and strictly proper scoring rules. Cost function market makers for complete markets over finite outcome spaces satisfying the properties of Abernethy et al. [2013] are known to be closely related to strictly proper scoring rules for eliciting probability distributions over the outcome space. In fact, each such cost function market maker is, in a strong technical sense, equivalent to a *market scoring rule* [Hanson 2003], a market in which each trader directly updates a market probability distribution and is paid for his improvement over the previous distribution according to a strictly proper scoring rule [Abernethy et al. 2013]. Similar ideas have been developed for incomplete markets with finite outcome and security spaces [Abernethy and Frongillo 2012], but to our knowledge, there is no prior work connecting cost function market makers for continuous outcome spaces with strictly proper scoring rules, despite existing work on proper scoring rules for continuous random variables [Matheson and Winkler 1976; Gneiting and Raftery 2007]. We formally make such a connection.

Other Related Work. Researchers have also designed cost function market makers with alternative desirable properties. By adapting the LMSR, Othman et al. [2010] introduced a market maker whose liquidity (quantified in terms of the rate at which prices react to trades) increases as trading volume increases. Othman and Sandholm [2011] generalized this intuition to provide a larger class of cost function market makers that have this liquidity sensitivity property, and Li and Wortman Vaughan [2013] gave a full axiomatic characterization of all cost function market makers with this and other desirable properties in a conjugate duality framework. Othman and Sandholm [2012] designed an alternative class of cost function market makers with adaptive liquidity based on the utility framework for market maker design [Chen and Pennock 2007]. All of these market makers are designed only for complete markets over finite outcome spaces, while our market makers apply to either complete or incomplete markets over any measurable outcome space.

This paper focuses on cost function market makers. It is worth noting that there are other market mechanisms, with different properties, designed for finite outcome spaces. For complete markets, Dynamic Parimutuel Markets [Pennock 2004; Mangold et al. 2005] also use a cost function to price securities. However the securities are parimutuel bets whose future payoff is not fixed a priori, but depends on the market activities. Brahma et al. [2012] and Das and Magdon-Ismail [2008] designed Bayesian learning market makers that maintain a belief distribution and update it based on the traders' behavior. Call markets have been studied to trade securities over combinatorial spaces [Fortnow et al. 2004; Chen et al. 2007; Agrawal et al. 2008; Ghodsi et al. 2008]. In a call market, participants submit limit orders and the market institution determines what orders to accept or reject.

Organization. We begin by formally defining a set of economic properties we would like a cost function for eliciting beliefs to satisfy in Section 2. We then elaborate on the mathematical duality that underlies our results in order to characterize these cost functions in Section 3. This characterization lets us construct cost functions for continuous random variables. In Section 4 we discuss the gap between theory and practice and offer a practical cost function for the interval $[0, 1]$. Finally, in Section 5 we connect our new characterization of cost functions to scoring rules.

2. THE ECONOMICS OF COST FUNCTIONS FOR ELICITING BELIEFS

In this section we review cost function market makers and describe four properties arguably desirable to elicit beliefs. These properties require that (1) the market admits a notion of prices, or equivalently, expectations, (2) these prices always reflect a feasible belief, (3) any trader can express his beliefs by trading in the market, and (4) any trader (myopically) maximizes his expected profit by expressing his beliefs. Together, these properties endow the marker maker with an implicit "market belief" that reflects information incorporated from traders, and incentivizes traders to incorporate this information.

2.1. Cost Function Market Makers

Let Ω be a set of *outcomes* or an *outcome space*, and assume that $(\Omega, \mathcal{F})$ is a measurable space. We might, for example, be interested in predicting which of two horses, Affirmed and Secretariat, will win a race, in which case $\Omega = \{\text{Affirmed wins, Secretariat wins}\}$. When Ω is finite, a natural $\sigma-$algebra $\mathcal{F}$ is the power set of Ω. Alternatively, we might be curious about the realized value of a continuous random variable like tomorrow's temperature or precipitation. These latter cases are represented with $\Omega = \mathbb{R}$ and $\mathcal{F}$ the Borel $\sigma-$algebra generated from the usual topology on the reals. We assume that the true outcome $\omega^* \in \Omega$

is privately selected by Nature at the start of time, and implicitly assume that traders have beliefs about this outcome that can be represented by probability measures.

A market lets traders purchase *portfolios* represented by *bounded measurable functions* $b : \Omega \to \mathbb{R}$. After the market closes, the outcome ω^* is revealed and the realized value of portfolio b is $b(\omega^*)$. Note that $b(\omega^*)$ can be negative, and traders can "sell" (or short sell) a portfolio b by purchasing $-b$.

Letting $\mathbf{B}$ be the set of all bounded measurable functions, a market offers a subset $B \subseteq \mathbf{B}$ of portfolios for purchase. We assume B is a subspace of $\mathbf{B}$. A market is said to be *complete* if $B = \mathbf{B}$ and *incomplete* otherwise.[1] One common method of generating the set B is to define it using a basis of bounded measurable functions (typically called *securities*). In this case a portfolio is a bundle of securities bought or sold together. In some cases, incomplete markets based on securities may be more tractable to run than complete markets. Our results apply to both types.

A *cost function market maker* is a special type of market maker that can be implemented using a potential function $C : B \to \mathbb{R}$ called the *cost function* that takes as input the market's current *liability* and returns a real. The market's current liability ℓ is the sum of purchased portfolios, which tells us the payments the market maker must make when different outcomes occur. We assume the market opens with an initial liability ℓ_0 (usually the constant zero function), and accepts a finite number of trades. If the market's current liability is ℓ and a trader purchases a portfolio b, the market's liability becomes $\ell' = \ell + b$ and the trader is charged $C(\ell') - C(\ell)$. When the outcome ω^* is revealed the net payoff to this trader is then

$$b(\omega^*) - \big(C(\ell + b) - C(\ell)\big) \;=\; (\ell' - \ell)(\omega^*) - \big(C(\ell') - C(\ell)\big) \qquad \text{(net payoff)}$$

where we use the notation $(\ell' - \ell)(\omega^*)$ to denote $\ell'(\omega^*) - \ell(\omega^*)$. Notice that the space of possible market liabilities is always the same as B, the space of available portfolios.

Cost function market makers have been studied for markets with finite outcome spaces. For example, the now-classic LMSR [Hanson 2003] belongs to this family. As mentioned in Section 1, Abernethy et al. [2011, 2013] characterized cost function market makers for finite outcome spaces and, in the case of infinite outcome spaces, when the market offers a basis of a finite number of securities. In their characterization, the potential function C is a function of *the number of shares of each security that have been purchased by all traders*. It is awkward to define C this way when a market may offer a basis of an infinite number of securities (e.g., securities that pay off \$1 if and only if $\omega^* \in (a, b)$ where $a \in [0, 1]$ and $b \in [0, 1]$). This necessitates our use of cost function which takes as an argument the liability function. For finite outcome spaces, a market offering a finite set of securities has a liability function $\ell(\omega) = \boldsymbol{\rho}(\omega) \cdot \mathbf{q}$ where $\mathbf{q}$ is the vector of purchased shares of each security and $\boldsymbol{\rho}(\omega)$ is the vector of payoffs of each security under outcome ω. When the market is complete and offers a set of Arrow-Debreu securities, one for each outcome and paying off \$1 if and only if the corresponding outcome happens, the liability function and the vector of purchased shares of each security are equivalent. Our cost functions can be applied for any measurable outcome space and include those for finite outcome spaces as special cases. Indeed, it can be shown that any market maker in the setting of Abernethy et al. [2013] can be written equivalently as a cost function based market maker in which the cost function takes as input the market's liability function, and in fact their framework is a special case of ours.

A cost function market maker enforces an arguably desirable property, *path independence*: the cost of purchasing a portfolio remains the same even if a trader splits the transaction into a number of consecutive transactions. Path independence implies that the amount

[1] Our concept of completeness is analogous to that used for markets with finite outcome spaces in that both require any contingent payoffs can be purchased in the market.

of money collected by the market maker is $C(\ell) - C(\ell_0)$, where ℓ is the current liability, regardless of the precise sequence of trades that led to this liability.

2.2. Market Prices/Expectations

The first of our four economic properties requires the market admit a notion of *price* for each portfolio. That is, for any current liability the market maker must have a well-defined instantaneous price for any portfolio, equal to the limit of the unit cost of purchasing the portfolio in ϵ portions at the current liability when $\epsilon \to 0$.

Property 1 (Existence of Market Prices). *At any liability $\ell \in B$, the market has a well-defined instantaneous price for any portfolio $b \in B$, equal to*

$$\nabla C(\ell; b) := \lim_{\tau \to 0} \frac{C(\ell + \tau b) - C(\ell)}{\tau}. \qquad \text{(market price)}$$

As described below, the price of a portfolio corresponds to the expected value of the portfolio according to the current "market belief." This allows traders to compare their subjective expectation of the value of a portfolio with the market expectation when making decisions.

2.3. Reasonable Prices

The next two properties let us extract a feasible belief from the market and ensure the market is capable of expressing the beliefs traders might have.

Defining these properties requires a notion of *agreement* between market prices and beliefs. Let $\mathcal{P}$ denote the set of probability measures over $(\Omega, \mathcal{F})$; if a trader has beliefs $p \in \mathcal{P}$ then his expectation for a portfolio b is $E_p[b] = \int_\Omega b \; \mathrm{d}p$. We say the market price *agrees with* a belief p if for all $b \in B$

$$\nabla C(\ell; b) = E_p[b] = \int_\Omega b \; \mathrm{d}p. \qquad \text{(market/belief agreement)}$$

We adopt the shorthand $\nabla C(\ell; \cdot) \cong_B p$ to signify this agreement.

Property 2 (Feasibility). *Every market price function $\nabla C(\ell; \cdot)$ agrees with a probability measure. That is, for each $\ell \in B$ there exists a $p \in \mathcal{P}$ such that $\nabla C(\ell; \cdot) \cong_B p$.*

Property 3 ($P-$Expressiveness). *Each probability measure in $P \subseteq \mathcal{P}$ agrees with a market price function on B. That is, for each $p \in P$ there exists an $\ell \in B$ such that $\nabla C(\ell; \cdot) \cong_B p$.*

Feasibility requires that the market's prices reflect *at least one* feasible belief. If the market is incomplete then its prices may reflect a set of feasible beliefs, and we refer to these sets (singleton or otherwise) as the market's *implicit beliefs*. Representing the beliefs of incomplete markets is discussed Section 3.1.

Feasibility is closely related to the no-arbitrage property of Abernethy et al. [2013], and is in fact equivalent for markets on finite outcome spaces, as discussed in Section 3.4.

If a market is $P-$expressive then traders with beliefs in P can change the market's implicit beliefs to match their own. Allowing P to be a subset of probability measures will be useful in Section 4 where we restrict attention to a practical subset, and will let us connect cost functions and scoring rules in Section 5.

2.4. Incentive Compatibility

Our fourth and final property requires that traders myopically maximize their expected profit for a trade by moving the market's prices to reflect their own.

Property 4 ((Myopic, Strict) $P-$Incentive Compatibility). *Assume $C : B \to \mathbb{R}$ is $P-$expressive. Then C is (myopic, strict) $P-$incentive compatible if for all $p \in P$, for* *all liabilities $\ell_p \in B$ such that $\nabla C(\ell_p;\cdot) \cong_B p$, for all $\ell, \ell' \in B$,*

$$\int_\Omega (\ell_p - \ell)\ \mathrm{d}p - \big(C(\ell_p) - C(\ell)\big) \ \geq\ \int_\Omega (\ell' - \ell)\ \mathrm{d}p - \big(C(\ell') - C(\ell)\big) \qquad \text{(incentive compatibility)}$$

with strict inequality when $C(\ell';\cdot) \ncong_B p$.

In other words, Property 4 guarantees a (myopic, risk neutral) trader has an incentive to "correct" the market any time the market's prices differ from his own beliefs.

$P-$incentive compatibility is analogous to the *information incorporation* property defined for markets with finite outcome spaces [Abernethy et al. 2013]. The information incorporation property requires that the instantaneous price of a bundle of securities weakly increases (or decreases) as a trader purchases (or sells) the bundle. Hence, if a trader's expected value of the bundle disagrees with the bundle's instantaneous price, the trader has a similar incentive to correct the market — he would find it profitable to buy or sell the bundle until the instantaneous price reflects his expected value. Together with the continuity of the cost function, which is assumed by Abernethy et al. [2013], the information incorporation property enforces the same convexity requirement on the cost function as incentive compatibility. We will prove this requirement for incentive compatibility in Theorem 1.

3. CHARACTERIZING COST FUNCTIONS FOR ELICITING BELIEFS

In this section we characterize cost functions for eliciting beliefs as the conjugates of a class of convex functions. This result and the connection between cost functions and scoring rules described in Section 5 rely on the duality between bounded measurable functions and probability measures introduced in Section 3.1. This introduction is technical out of necessity, but crucial to understanding our paper's contribution.

3.1. The Duality of Market Liabilities and Beliefs: Mathematical Background

The duality between the bounded measurable functions (representing both portfolios and market liabilities) and probability measures (representing beliefs) is critical to our perspective. This subsection offers a brief introduction to the aspects of this duality necessary for understanding our theorem statements and supporting analysis. It details our refinement of the notions of subdifferentials and strict convexity, as well as two useful facts that are used in proofs throughout the paper. The first is the well-known "conjugate-subgradient" theorem and the latter is a collection of equivalences relating strict convexity and the subdifferential of a convex function.

A *Banach space* is a normed vector space that is complete[2] with respect to its norm. Every Banach space $\mathbf{X}$ admits a *topological* or *continuous dual space* $\mathbf{Y}$ of all continuous[3] linear functions $y : \mathbf{X} \to \mathbb{R}$. This dual space is also a Banach space with pointwise addition and scalar multiplication for functions and the *dual norm*

$$||y|| := \sup_{x \in \mathbf{X}, ||x|| \leq 1} y(x). \qquad \text{(dual norm)}$$

A space and its dual also admit a natural bilinear form $\langle \cdot, \cdot \rangle : \mathbf{X} \times \mathbf{Y} \to \mathbb{R}$ that is linear in both arguments and defined as $\langle x, y \rangle := y(x)$.

Let $(\Omega, \mathcal{F})$ be a measurable space and $\mathbf{B}$ the set of all bounded measurable functions[4] $b : \Omega \to \mathbb{R}$. $\mathbf{B}$ is a Banach space when equipped with the supremum norm, $||b|| = \sup_{\omega \in \Omega} b(\omega)$,

[2]A space is complete if every Cauchy sequence has its limit inside the space.

[3]When we say continuous or lower semicontinuous, it will always be with respect to the norm topology.

[4]Here $\mathbb{R}$ is endowed with the usual Borel topology. Note that while we require each function be bounded (so the supremum norm is defined) this does not mean the set of functions is bounded.

and the convention of pointwise addition and scalar multiplication. Its dual space contains all countably additive (and finite) measures, and the bilinear form between the countably additive measures and bounded measurable functions is the Lebesgue integral

$$\langle b, \mu \rangle = \mu(b) = \int_\Omega b \, \mathrm{d}\mu. \qquad \text{(bilinear form)}$$

The set of probability measures $\mathcal{P}$ is a convex subset of the Banach space of countably additive measures, and thus a subset of the dual space of $\mathbf{B}$.

A cost function $C : B \to \mathbb{R}$ satisfying our four economic properties relates its liabilities to beliefs through market prices, and the structure of these beliefs depends on the dual space of offered portfolios. By definition, each element of the dual space of B is simply a linear map. We will see (in Corollary 1) that a cost function associates each liability with an element of this dual space, and this element represents the market's implicit belief. If the market is complete ($B = \mathbf{B}$) then this dual space simply contains $\mathcal{P}$ and so the market's price agrees with only one probability measure (i.e., for any liability ℓ, $\nabla C(\ell, \cdot) \cong_B p$ for a single measure p). If the market is incomplete ($B \subset \mathbf{B}$), however, then the dual space of B may be *coarser* than that of $\mathbf{B}$ and its elements can represent sets of measures. That is, an incomplete market may have prices consistent with more than one probability measure.

We'll use p_B to denote the element of the dual space satisfying $\langle b, p_B \rangle = \langle b, p \rangle$ for all $b \in B$, so $p_B \cong_B p$. Note that if $p_B = p'_B$, then $E_p[b] = E_{p'}[b]$ for all $b \in B$. We'll use P_B to denote the set $\{p_B | p \in P\}$ and $\mathcal{P}_B$ to denote the set $\{p_B | p \in \mathcal{P}\}$. To reiterate, there may be a many-to-one mapping from probability measures to the market maker's implicit beliefs since these implicit beliefs may fail to distinguish between two or more probability measures; p_B is how we denote the image of a probability measure p under this mapping.

Convex functions describe a class of relationships between spaces in duality via the notion of subdifferentials. We'll see that it is precisely these relationships that encode our desired pairing between market liabilities and beliefs. Formally, letting $\mathbf{X}$ be a Banach space, $\mathbf{Y}$ its dual space, and $\bar{\mathbb{R}} = [-\infty, \infty]$ the extended reals, the subdifferential of a convex function $f : \mathbf{X} \to \bar{\mathbb{R}}$ is the function

$$\partial f(x_0) = \{y \in \mathbf{Y} | f(x_0) - f(x_1) \leq \langle x_0 - x_1, y \rangle, \forall x_1 \in \mathbf{X}\} \qquad \text{(subdifferential)}$$

mapping points in $\mathbf{X}$ to sets of elements in $\mathbf{Y}$ satisfying the subdifferential inequality. We'll say a convex function is *proper* if it is nowhere negative infinity and somewhere real-valued.

The dual space of the countably additive measures contains the bounded measurable functions $\mathbf{B}$ *and other functions we will not be interested in*, so we introduce a refinement of the subdifferential that excludes the latter. Letting $Y \subseteq \mathbf{Y}$, the $Y-$subdifferential of $f : \mathbf{X} \to \mathbb{R}$ is the function

$$\partial_Y f(x_0) = \partial f(x_0) \cap Y. \qquad (Y\text{−subdifferential})$$

We say that a function f is $Y-$*subdifferentiable* at a point x if $\partial_Y f(x)$ is non-empty. If the function's $Y-$subdifferential is nowhere empty we simply say it is $Y-$subdifferentiable. Let $\mathrm{dom}(\partial_Y f)$ denote the subset of $\mathbf{X}$ at which f is $Y-$subdifferentiable. We say a function's subdifferential contains Y if each element of Y is part of the subdifferential at some point. A function f has *disjoint* $Y-$*subdifferentials* when

$$\partial_Y f(x_0) \cap \partial_Y f(x_1) = \emptyset, \ \forall x_0 \neq x_1 \in \mathbf{X} \qquad \text{(disjoint subdifferentials)}$$

and is *strictly convex where* $Y-$*subdifferentiable* when

$$\alpha f(x_0) + (1 - \alpha) f(x_1) > f(\alpha x_0 + (1 - \alpha) x_1), \quad \text{(strictly convex where subdifferentiable)}$$

for all $\alpha \in (0, 1)$ and $x_0, x_1 \in \mathbf{X}$ such that $x_0, x_1, \alpha x_0 + (1 - \alpha) x_1 \in \mathrm{dom}(\partial_Y f)$. In other words, a function is strictly convex where $Y-$ subdifferentiable if the convex inequality holds strictly whenever f is $Y-$subdifferentiable at the three points in question.

To describe the subdifferential we also need the idea of a convex conjugate. The convex conjugate of a function $f : \mathbf{X} \to \bar{\mathbb{R}}$ is the function $f^* : \mathbf{Y} \to \bar{\mathbb{R}}$ defined as

$$f^*(y) = \sup_{x \in \mathbf{X}} \ \langle x, y \rangle - f(x) \tag{conjugate}$$

and its relationship with the subdifferential is described by the conjugate-subgradient theorem, our statement of which is adopted from Barbu and Precupanu [2012].

Fact 1 (Conjugate-Subgradient Theorem). *Let $\mathbf{X}$ be a Banach space, $\mathbf{Y}$ its topological dual space and $f : \mathbf{X} \to \bar{\mathbb{R}}$ a proper convex and lower semicontinuous function. Then the following four properties are equivalent:*

(1) $y \in \partial f(x)$
(2) $f(x) + f^*(y) \leq \langle x, y \rangle$
(3) $f(x) + f^*(y) = \langle x, y \rangle$
(4) $x \in \partial f^*(y)$.

The conjugate of a proper convex function is also proper and lower semicontinuous. We'll use the fact that the biconjugate of a proper and lower semicontinuous function agrees with the original function where the original is defined. That is, for all our purposes $f = f^{**}$ when f is a proper and lower semicontinuous convex function. (The biconjugate is the lower semicontinuous closure of f.)

Finally, a function $f : \mathbf{X} \to \bar{\mathbb{R}}$ is *Gâteaux differentiable* at an algebraic interior point[5] $x \in \mathbf{X}$ if the function

$$\nabla f(x; h) = \lim_{\tau \to 0} \frac{f(x + \tau h) - f(x)}{\tau} \tag{Gâteaux differential}$$

is well-defined (i.e., the limit exists for all $h \in \mathbf{X}$) and if this function is a continuous linear function of h (i.e., the function is an element of $\mathbf{Y}$). If a convex function is Gâteaux differentiable at a point, then its subdifferential is a singleton consisting of its Gâteaux differential there. Conversely if a convex function is finite and continuous at a point and its subdifferential is a singleton there then the function is Gâteaux differentiable at that point and that subdifferential is its Gâteaux differential ([Barbu and Precupanu 2012], p.87).

We conclude with a collection of equivalences relating the subdifferential to strict convexity. These equivalences appear to be folk knowledge and we offer a discussion of their provenance and an explicit proof in the appendix.[6]

Fact 2 (Disjoint Subdifferential Equivalences). *Let $\mathbf{X}$ be a Banach space, $\mathbf{Y}$ its dual space, $Y \subseteq \mathbf{Y}$, and $f : \mathbf{X} \to \bar{\mathbb{R}}$ a proper convex and lower semicontinuous function. Then the following are equivalent:*

(1) f has disjoint $Y-$subdifferentials
(2) f is strictly convex where $Y-$subdifferentiable
(3) the subgradient inequality, $f(x) - f(x') \leq \langle x - x', y \rangle$, $\forall x' \in \mathbf{X}$, holds strictly whenever $x \neq x'$ for all x and $y \in \partial_Y f(x)$
(4) the subdifferentials of f^* on the set Y consist of singleton sets.

3.2. A Dual Characterization

We can now use this duality to characterize cost functions for eliciting beliefs as the conjugates of a class of convex functions, yielding our main characterization result.

[5]An algebraic interior point of a set X is a point of X such that every line through that point lies in the affine hull of X. The functions we are interested in applying this differential to will always be defined over entire Banach spaces, and so every point in their domain will (trivially) be an algebraic interior point.
[6]An appendix with omitted proofs appears in the version of this paper available on the authors' websites.

Theorem 1 (Cost Functions for Eliciting Beliefs Characterization). *Let $B \subseteq \mathbf{B}$ be a Banach space and $P \subseteq \mathcal{P}$ a set of probability measures. A cost function $C : B \to \mathbb{R}$ satisfies the existence of market prices, feasibility, $P-$expressiveness, and $P-$incentive compatibility if and only if it can be written as*

$$C(\ell) = \sup_{p_B \in \mathcal{P}_B} \langle \ell, p_B \rangle - R(p_B)$$

where $R : \mathcal{P}_B \to \bar{\mathbb{R}}$, is a proper lower semicontinuous and convex function whose subdifferential contains B with $P_B \subseteq \operatorname{dom}(\partial_B R) \subseteq \mathcal{P}_B$ and that is strictly convex where $B-$subdifferentiable.

Proof. We begin by showing the conjugate of such a function R satisfies Properties 1–4, then conclude by arguing the necessity of such a function existing.

Part 1: Sufficiency. Let R be a function meeting the statement's criteria, and C the conjugate of R. Since C is the conjugate of a convex function it is also proper, lower semicontinuous, and convex.

We first show C satisfies Property 1 by being Gâteaux differentiable. This first requires showing C is real-valued and continuous. C is proper and so it never attains $-\infty$, so we only need to show it never attains $+\infty$. We can bound $|\langle \ell, p \rangle| \leq ||\ell|| = \sup_{\omega \in \Omega} |\ell(\omega)|$ since p is a probability measure and ℓ is a bounded measurable function. Thus if C is somewhere $+\infty$ then R is unbounded below, contradicting its being proper. We conclude that C is real-valued and since it is real-valued and lower semicontinuous on B, a Banach space, it is continuous ([Barbu and Precupanu 2012], p. 74).

Since C is real-valued and continuous it is subdifferentiable everywhere ([Barbu and Precupanu 2012], p.85, [Minty 1964]). Since we assumed R was strictly convex where $B-$subdifferentiable, Fact 2 says C has singleton sets for its subdifferentials (since its domain is B). This implies C is Gâteaux differentiable, and this says the limit

$$\nabla C(\ell, b) = \lim_{\tau \to 0} \frac{C(\ell + \tau b) - C(b)}{\tau}$$

is well-defined for all $b \in B$, the same limit required by Property 1.

Properties 2 and 3 follow directly from our description of the subdifferential. Since we assumed R was $B-$subdifferentiable on a subset of $\mathcal{P}_B$ and that its subdifferential contains B, each Gâteaux differential of C agrees with a probability measure. We also assumed R was $B-$subdifferentiable on a superset of P_B, so for any $p \in P$ there exists a Gâteaux differential of C that agrees with p.

Finally, we conclude by demonstrating Property 4; letting $\ell_p \in B$ be any liability such that $\nabla C(\ell_p; \cdot) \cong_B p$ for some p with $p_B \in P_B$, this property requires $\langle \ell_p - \ell, p \rangle - \big(C(\ell_p) - C(\ell)\big) \geq \langle \ell' - \ell, p \rangle - \big(C(\ell') - C(\ell)\big)$, for all $\ell, \ell' \in B$, with strict inequality when $\nabla C(\ell'; \cdot) \ncong_B p$. We can rearrange this into a subgradient inequality $C(\ell_p) - C(\ell') \leq \langle \ell_p - \ell', p \rangle$ that is strict when $\nabla C(\ell'; \cdot) \ncong_B p$. If this holds with equality, i.e., $\langle \ell', p \rangle - C(\ell') = \langle \ell_p, p \rangle - C(\ell_p)$, Fact 1 tell us this implies p_B is a subgradient of C at ℓ', too, and so $\nabla C(\ell'; \cdot) \cong_B p$ and Property 4 is satisfied.

Part 2: Necessity. Let C be a cost function satisfying Properties 1–4. We first show that C is proper, lower semicontinuous, and convex by demonstrating it can be expressed as the pointwise supremum of a family of continuous affine functions. We then use these properties to argue about the conjugate of C, called R.

Let P be the maximal set of probability measures for which C is $P-$expressive, and let L_p be the set of liability functions ℓ such that $C(\ell; \cdot) \cong_B p$. We define a family of continuous affine functions $f_p(\ell) = \langle \ell - \ell_p, p \rangle + C(\ell_p)$ were ℓ_p is any member of L_p, with p ranging over P. Note that the choice of $\ell_p \in L_p$ is immaterial to the definition of these functions since

Property 4 implies $\langle \ell - \ell_p, p\rangle + C(\ell_p) = \langle \ell - \ell'_p, p\rangle + C(\ell'_p)$ for all $\ell_p, \ell'_p \in L_p$. The pointwise supremum of this family of functions is defined as

$$S(\ell) \;=\; \sup_{p\in P}\; f_p(\ell) \;=\; \sup_{p\in P}\; \langle \ell - \ell_p, p\rangle + C(\ell_p).$$

Note that Properties 1 and 2 together imply each $\ell \in B$ is associated with at least one probability measure. That is, for each $\ell \in B$ there exists $p \in P$ such that $\ell \in L_p$. This implies that $S(\ell) \geq C(\ell)$ for all ℓ. In fact, S must agree with C on its domain, since if this were not the case, then there would exist an $\ell \in B$ such that $\sup_{p\in P}\ \langle \ell - \ell_p, p\rangle + C(\ell_p) > C(\ell)$, which would violate Property 4. So we conclude C is a pointwise supremum of a family of continuous affine functions. Since C is real-valued (by definition) it is proper, and if a proper function is also the pointwise supremum of a family of continuous affine functions it is lower semicontinuous and convex ([Barbu and Precupanu 2012], p.80).

Since C is proper, lower semicontinuous, and convex, it has a conjugate function R and is equal to its biconjugate, i.e., $C = R^*$. Immediately, we know R is also proper, lower semicontinuous, and convex. Thus C is the conjugate function of R, and to complete this portion of the proof we must show that:

(1) R has a subdifferential containing B,
(2) R is subdifferential on a superset of P_B and subset of $\mathcal{P}_B$, and
(3) R is strictly convex where $B-$subdifferentiable.

(Note that we wrote $R : \mathcal{P}_B \to \mathbb{R}$, although R may be well-defined beyond the set $\mathcal{P}_B$, we restrict attention to this set since the supremum expression defining the cost function is restricted to it as a result of the feasibility property it satisfies.)

Properties 1 and 2 imply C is Gâteaux differentiable, and thus it is subdifferentiable and has singleton sets for its subdifferential. The conjugate-subgradient theorem tells us this implies B is part of the subdifferential of R, and these subgradients of R appear on the set $P_B \subseteq \mathcal{P}_B$, since P_B describes the subdifferential of C. Applying Fact 2 also tell us that R is strictly convex where $B-$subdifferentiable, and this concludes the proof. $\square$

As a corollary we can describe the cost function directly.

Corollary 1 (Describing Cost Functions for Eliciting Beliefs). *Let $B \subseteq \mathbf{B}$ be a Banach space and $P \subseteq \mathcal{P}$ a set of probability measures. A cost function $C : B \to \mathbb{R}$ satisfies the existence of market prices, feasibility, $P-$expressiveness, and $P-$incentive compatibility if and only if it is (real-valued) continuous, convex, and Gâteaux differentiable, and its subdifferential contains P_B and is a subset of $\mathcal{P}_B$.*

Theorem 1 says a cost function for eliciting beliefs c can be expressed as the conjugate of a function $R : \mathcal{P}_B \to \bar{\mathbb{R}}$ from a particular class of convex functions. Intuitively, this expression says the cost function determines its value by assuming a belief that maximizes the expected value of its liabilities subject to a "regularization" term R that adds stability, analogous to techniques applied in machine learning. It is the choice of this regularization function that identifies a cost function for eliciting beliefs.

This is the same intuition expressed in Chen and Wortman Vaughan [2010] and Abernethy et al. [2013] for the special case of a finite set of outcomes. Abernethy et al. [2013] also showed this dual space can be represented as a subset of finite dimensional Euclidean space, the convex hull of security payoffs. In Section 4 we will apply our more general characterization to infinite outcome spaces that do not allow this nicety.

3.3. Markets with Bounded Loss

A practical and desirable property of a market not directly related to or necessary for eliciting beliefs is that it has bounded worst-case loss.

Property 5 (Bounded Worst-Case Loss). *Assume C opens with liability ℓ_0, the constant zero function. Then there exists some $k \in \mathbb{R}$ such that*

$$\sup_{\ell \in B, \omega \in \Omega} \ell(\omega) - (C(\ell) - C(\ell_0)) \leq k. \qquad \text{(bounded worst-case loss)}$$

It turns out that this is equivalent to the conjugate function R being bounded and expressive enough to allow any belief to be arbitrarily approximated. This generalizes the result of Abernethy et al. [2013] that states that markets for finite outcome spaces have bounded loss when R is bounded.

Theorem 2 (Markets with Bounded Loss). *Let C be a cost function satisfying the conditions in the statement of Theorem 1, with $P \subseteq \mathcal{P}$ the maximal set for which C is $P-$expressive. Then C has bounded worst-case loss if and only if its conjugate R is bounded on P_B and P_B is dense in $\mathcal{P}_B$. If this is the case, then the loss is bounded by $\sup_{p_B \in P_B} R(p_B) - \inf_{p_B \in P_B} R(p_B)$.*

Proof. First assume that the conjugate R of C is bounded on P_B and that P_B is dense in $\mathcal{P}_B$. Let $\delta(\omega)$ be the Dirac measure assigning probability to the single state $\omega \in \Omega$. Expanding out C using the characterization in Theorem 1, the worst-case loss is

$$\sup_{\ell \in B, \omega \in \Omega} \langle \ell, \delta(\omega) \rangle - \sup_{p_B \in P_B} \{\langle \ell, p_B \rangle - R(p_B)\} + \sup_{p_B \in P_B} \{\langle \ell_0, p_B \rangle - R(p_B)\}. \tag{1}$$

Note that we have replaced $\mathcal{P}_B$ with P_B in the expansion of C; this is because P is the maximal set for which C is $P-$expressive and so by the conjugate-subgradient theorem, $P_B = \mathrm{dom}(\partial_B R)$ and the supremum must be achieved at a point in this set. Since $\ell_0 = 0$, the last supremum can be replaced with $-\inf_{p_B \in P_B} R(p_B)$. Now, let $\epsilon > 0$ be an arbitrarily small real number. Since we assumed P_B is dense in $\mathcal{P}_B$, for any ℓ there exists $p_B^{\ell,\epsilon} \in P_B$ such that $\sup_{\omega \in \Omega} \langle \ell, \delta(\omega) - p_B^{\ell,\epsilon} \rangle \ \leq \ \epsilon$. Thus we can upper bound Equation 1 by

$$\begin{aligned} &\sup_{\ell \in B, \omega \in \Omega} \langle \ell, \delta(\omega) \rangle - \{\langle \ell, p_B^{\ell,\epsilon} \rangle - R(p_B^{\ell,\epsilon})\} - \inf_{p_B \in P_B} R(p_B) \\ &\qquad \leq \epsilon + \sup_{\ell \in B} R(p_B^{\ell,\epsilon}) - \inf_{p_B \in P_B} R(p_B) \leq \epsilon + \sup_{p_B \in P_B} R(p_B) - \inf_{p_B \in P_B} R(p_B). \end{aligned}$$

Since ϵ can be set arbitrarily close to 0, this shows that the worst-case loss is bounded by $\sup_{p_B \in P_B} R(p_B) - \inf_{p_B \in P_B} R(p_B)$ as desired.

Next consider the case in which R is unbounded on P_B. We already know R is bounded below following the argument made in the proof of Theorem 1, so we assume R is unbounded above. Since the integral of ℓ with respect to a probability measure is always weakly less than its supremum norm, we have that for any $\ell \in B$ and $p_B \in P_B$, $\sup_{\ell \in B} \ ||\ell|| - \langle \ell, p_B \rangle \geq 0$. Since R is unbounded on P_B, this implies that for any $M \in \mathbb{R}$, there exists a $p_B \in B$ such that $R(p_B) \geq M$ and subsequently, for all ℓ, $\sup_{\ell \in B} \ ||\ell|| - \langle \ell, p_B \rangle + R(p_B) \ \geq \ M$. Since $P_B = \mathrm{dom}(\partial_B R)$, there exists some ℓ such that p_B is in the subdifferential of R at ℓ, and for this ℓ, $C(\ell) = \langle \ell, p_B \rangle - R(p_B)$. Therefore, for this particular ℓ, $||\ell|| - C(\ell) \geq M$, and so

$$\sup_{\ell \in B, \omega \in \Omega} \ell(\omega) - (C(\ell) - C(\ell_0)) \geq M + C(\ell_0).$$

Since this argument holds for arbitrarily large M and $C(\ell_0)$ is a constant, this implies the market maker has unbounded worst-case loss.

Finally, assume R is bounded on P_B, but P_B is not dense in $\mathcal{P}_B$. This implies there exists $\ell \in B$ and $p'_B \in \mathcal{P}_B$ such that $\inf_{p_B \in P_B} \langle \ell, p'_B - p_B \rangle \geq \epsilon$. for some $\epsilon > 0$, which implies that

$$||\ell|| - \sup_{p_B \in P_B} \langle \ell, p_B \rangle \geq \langle \ell, p'_B \rangle - \sup_{p_B \in P_B} \langle \ell, p_B \rangle = \inf_{p_B \in P_B} \langle \ell, p'_B - p_B \rangle \geq \epsilon.$$

Letting R be bounded below by r, we have that for any $k > 0$,

$$\sup_{\omega\in\Omega} k\ell(\omega) - (C(k\ell) - C(\ell_0)) = ||k\ell|| - \sup_{p_B \in P_B} \{\langle k\ell, p_B\rangle - R(p_B)\} + C(\ell_0) \geq k\epsilon + r + C(\ell_0).$$

Since k can be made arbitrarily large while r and $C(\ell_0)$ are constants, this implies the worst-case loss is unbounded if P_B is not dense in $\mathcal{P}_B$. $\square$

3.4. A Lack of Arbitrage

Instead of feasibility, Abernethy et al. [2013] required that markets do not allow arbitrage, i.e., a trader can never purchase a portfolio with a guaranteed positive net payoff regardless of the outcome. Formally, a cost function C permits no arbitrage if

$$C(\ell + b) - C(\ell) \geq \inf_{\omega\in\Omega} b(\omega) \qquad \text{(no arbitrage)}$$

for all $\ell, b \in B$. With finite outcome spaces, feasibility and no arbitrage are equivalent. With infinite outcomes spaces, they may differ. However, a cost function satisfying Properties 1–4 automatically satisfies this property.

Theorem 3 (No Arbitrage). *Let $B \subseteq \mathbf{B}$ be a Banach space and $P \subseteq \mathcal{P}$ a set of probability measures. A cost function $C : B \to \mathbb{R}$ satisfying existence of market prices, feasibility, $P-$expressiveness, and $P-$incentive compatibility also permits no arbitrage.*

Proof. Corollary 1 tells us the cost function is convex and Gâteaux differentiable. We apply the subgradient inequality, $C(\ell + b) - C(\ell) \geq \langle b, p_B\rangle$, where p_B is the subgradient of C at ℓ. Since $\langle b, p_B\rangle \geq \inf_{\omega\in\Omega} b(\omega)$, the no-arbitrage inequality holds. $\square$

4. COST FUNCTIONS FOR [0, 1]

Predicting the outcome of a continuous random variable, such as tomorrow's precipitation, is a natural problem. However, prior efforts have had no success designing even a single market maker that can support betting on the outcome of a continuous random variable without requiring that the market maker incur an infinite loss in the worst case [Gao et al. 2009; Gao and Chen 2010]. Our framework enables us create bounded loss cost functions for continuous outcome spaces like the [0, 1] interval. However, a straightforward construction of these cost functions reveals an interesting gap between theory and practice, in that cost functions satisfying Properties 1–5 can exhibit odd and undesirable behavior. We begin this section by describing this odd behavior, and then suggest a practical solution. This topic could easily be the subject of its own paper; here we give only one example of a practical solution as an illustration of how our theory can be applied.

4.1. Biased Cost Functions

Consider a finite outcome space $\Omega = \{\omega_1, \cdots, \omega_n\}$. An example of a cost function satisfying Properties 1–5 for this outcome space is

$$C(\ell) = \sup_{p\in\mathcal{P}} \sum_{i=1}^{n} \ell(\omega_i)p(\omega_i) - \sum_{i=1}^{n} p^2(\omega_i),$$

which uses a quadratic function R. This cost function exhibits a nice *anonymity* or *unbiasedness* property in that whenever $\ell(\omega_i) = \ell(\omega_j)$ for any two outcomes, the supremum (and therefore, market price vector) is attained at a probability distribution that assigns equal weight to both outcomes.

Now assume Ω is ordered and let F_p be the cumulative distribution function (CDF) associated with p, so $F_p(\omega_1) = p(\omega_1)$, $F_p(\omega_2) = p(\omega_1) + p(\omega_2)$, and so on. Another cost function also satisfying Properties 1–5 is

$$C(\ell) = \sup_{p \in \mathcal{P}} \sum_{i=1}^{n} \ell(\omega_i) p(\omega_i) - \sum_{i=1}^{n} F_p^2(\omega_i).$$

This cost function is not so nice. If $\ell = 0$ then the market's implicit belief would assign probability one to to a single outcome. In short, markets like this appear *biased*.

Intuitively, conjugates defined in terms of probability density functions (PDFs) tend to be unbiased because they operate on local structure, while those defined on CDFs appear biased. While beliefs in every finite outcome space can be represented as a PDF, the same is not true with infinite outcome spaces. The natural Borel probability measures on the [0, 1] interval, for example, correspond with the set of CDFs. This creates a unique practical challenge for creating a cost function market maker for the interval, since the first cost function doesn't generalize, and we are unlikely to be satisfied with the generalization of the second, (i.e., $C(\ell) = \sup_{p \in \mathcal{P}} \int_\Omega \ell \; \mathrm{d}p - \int_\Omega F_p^2 \; \mathrm{d}\lambda$) since it maps uniform liability to measures with probability 1 on a single outcome.

4.2. Predictions on [0, 1]

Before describing more practical cost functions we will need a detour to set up some results and notation. When discussing the [0, 1] interval we associate it with the measurable space $([0, 1], \mathcal{B})$ where $\mathcal{B}$ is the Borel $\sigma-$algebra generated from the usual (subspace) topology on the interval. The probability measures $\mathcal{P}$ on this interval correspond with CDFs, and this correspondence is such that a (strictly) convex function of one is also a (strictly) convex function of the other (*mutatis mutandi*).

Let λ denote Lebesgue measure; every measure $p \in \mathcal{P}$ can be decomposed into three parts with respect to λ: a *pure point* part p_{pp} consisting of a countable number of atoms, an *absolutely continuous* part p_{cont} that admits a density function, and a *singular continuous* part p_{sing} that is continuous and does not admit a density function[7]. We'll let $\mathcal{P}_{cont}$ be the set of all probability measures that are absolutely continuous with respect to λ and $\mathcal{P}_{practical}$ be probability measures consisting of only pure point and absolutely continuous parts. We will define our cost functions in terms of strictly convex functions of probability measures. The following result lets us define a broad class of them.

Lemma 1 (Strictly Convex Functions of Absolutely Continuous Measures). *Let $\psi : \mathbb{R} \to \mathbb{R}$ be a strictly convex function. Then the function*[8]

$$\Psi(\mu) = \int_0^1 \psi\left(\frac{\mathrm{d}\mu}{\mathrm{d}\lambda}\right) \; \mathrm{d}\lambda$$

is a strictly convex function of measures μ absolutely continuous with respect to the Lebesgue measure λ.

4.3. A Practical Cost Function

Our practical cost function based market will offer a security for every subinterval of $[0, 1]$, worth \$1 if the outcome is in the subinterval and \$0 otherwise. The space of portfolios B generated from this basis is not a Banach space but the space of *simple functions* B_{simple}.[9] This will prevent us from immediately applying our previous results and suggests additional refinements may be necessary for some applications.

---

[7] The canonical example of a singular continuous measure is the devil's staircase or Cantor distribution that has uniform support on the Cantor set.

[8] We use the expression $\frac{\mathrm{d}\mu}{\mathrm{d}\lambda}$ to stand for the Radon-Nikodym derivative of μ with respect to λ.

[9] A simple function is any function that attains only a finite number of values on the interval.

Following the discussion that opened this section, we might restrict the market maker to beliefs in $\mathcal{P}_{cont}$ and use the cost function

$$C(\ell) = \sup_{p \in \mathcal{P}_{cont}} \int_\Omega \ell \; \mathrm{d}p - \int_\Omega \left(\frac{\mathrm{d}p}{\mathrm{d}\lambda} \right)^2 \; \mathrm{d}\lambda$$

While this cost function is unbiased, it never prices single points positively. This lets traders purchase any countable set of points for free. Instead we adopt a "hybrid" cost function that allows beliefs in $\mathcal{P}_{practical}$; restricting traders to beliefs in this set loses little generality.

To formally state that this market maker is unbiased, we need a formal definition of unbiasedness. We first state our definition, and then describe the intuition behind it.

PROPERTY 6 (LEBESGUE UNBIASED). *Let $B \subseteq \mathbf{B}$ and $P \subseteq \mathcal{P}$. A cost function $C : B \to \mathbb{R}$ is Lebesgue unbiased on B and P if for all $p \in P$ such that $\nabla C(\ell; \cdot) \cong_B p$, for all $\Omega_0, \Omega_1 \in \mathcal{B}$, if $\ell(\omega_0) = \ell(\omega_1)$ for all $\omega_0 \in \Omega_0$ and $\omega_1 \in \Omega_1$, then*

(1) if $\lambda(\Omega_0) = \lambda(\Omega_1) = 0$, then $p(\Omega_0) = p(\Omega_1)$, and
(2) if $\lambda(\Omega_0), \lambda(\Omega_1) > 0$, then $p(\Omega_0)/\lambda(\Omega_0) = p(\Omega_1)/\lambda(\Omega_1)$.

In short, Property 6 says that if two measurable sets have the same liability everywhere, then the market's beliefs must assign them likelihood in proportion to their Lebesgue measures. For example, if both $[.2, .3]$ and $(.4, .6)$ had the same associated liabilities then the market's implicit belief would assign $[.2, .3]$ half the likelihood of $(.4, .6)$ since it is half the "size" of the latter with respect to the Lebesgue measure. Defining our notion of bias with respect to a reference measure is essential to its being well-defined. Even when Ω is finite our intuition could be formalized as being unbiased with respect to the counting measure. This also highlights that the choice of reference is dependent on the setting and our tastes.

We will see that one especially important feature of a cost function's being unbiased is that it *can be solved using a convex program with a finite number of variables whenever the liability is simple*. This suggests the property is computationally desirable too.

We now formally define our proposed practical cost function and show it is unbiased. This cost function penalizes point masses and "rewards" more continuous beliefs. This mixed structure is important. Imagine a constant liability function; the market's corresponding beliefs minimize its penalty or regularization function alone. If, however, continuous beliefs were not rewarded, then the market could create arbitrarily small point masses in its beliefs, essentially taking a penalty of zero, instead of forming continuous beliefs. Rewarding continuous beliefs encourages the market to hold them as an alternative to these ever smaller collections of point masses. Note, however, that if the liability function "spikes" at a single point, the market has no alternative and will (if the spike is high enough) decide to create a point mass there. The use of the negative arctan function in particular is not crucial; what is important is that it is a bounded and decreasing strictly convex function.

When stating this cost function we abuse notation and let p_{pp} also stand for the (countable) set where it has support.

THEOREM 4 (PRACTICAL COST FUNCTION MARKET MAKER). *The cost function*

$$C(\ell) = \sup_{p \in \mathcal{P}_{practical}} \int_\Omega \ell \; \mathrm{d}p - \left(\sum_{\omega \in p_{pp}} p(\omega)^2 - \int_\Omega \arctan \left(\frac{\mathrm{d}p}{\mathrm{d}\lambda} \right) \; \mathrm{d}\lambda \right) \tag{2}$$

is Lebesgue unbiased on B_{simple} and $\mathcal{P}_{practical}$.

Additionally, this cost function can be calculated in time linear in the number of securities purchased so far. If we let the simple function ℓ be described as a collection of tuples $(\Omega_0, k_0), (\Omega_1, k_1), \ldots, (\Omega_n, k_n)$ consisting of measurable sets and the reals to which ℓ maps them, I be an index set over all such tuples, and J an index set over tuples where $\lambda(\Omega_i) = 0$, we can solve for this cost function using the convex program

$$\max \sum_{i \in I} k_i p_i - \left(\sum_{j \in J} p_j^2 - \sum_{h \in I - J} \arctan\left(\frac{p_h}{\lambda(\Omega_h)} \right) \right) \quad \text{s.t.} \quad \sum_{i \in I} p_i = 1$$

when it has simple liabilities. Further, this convex program admits a unique solution.

This market falls outside of our framework since B_{simple} is not a Banach space. However, variants of the desirable economic properties described in the previous sections hold, and the market has bounded worst-case loss, as the following theorem shows.

THEOREM 5 (BOUNDED LOSS). *The worst-case loss of the market maker specified by the cost function in Equation 2 with $\ell_0 = 0$ is bounded by $1 + \pi/4 \leq 1.79$.*

PROOF. Worst-case loss is defined as $\sup_{\ell \in B_{simple}, \omega \in \Omega} \ell(\omega) - \big(C(\ell) - C(\ell_0)\big)$. Expanding the definition of C, we have

$$\begin{aligned} C(\ell_0) &= \sup_{p \in \mathcal{P}_{practical}} 0 - \sum_{\omega \in p_{pp}} p(\omega)^2 + \int_\Omega \arctan\left(\frac{\mathrm{d}p}{\mathrm{d}\lambda}\right) \mathrm{d}\lambda && (\ell_0 = 0) \\ &= \sup_{p \in \mathcal{P}_{practical}} \int_\Omega \arctan\left(\frac{\mathrm{d}p}{\mathrm{d}\lambda}\right) \mathrm{d}\lambda && (\textstyle\sum_{\omega \in p_{pp}} p(\omega)^2 \text{ is always positive}) \\ &= \arctan(1) = \pi/4. && (\text{the cost function is Lebesgue unbiased}) \end{aligned}$$

For any ℓ, let $\omega \in \Omega$ be the outcome at which ℓ attains its supremum, and let $\delta(\omega)$ be the Dirac measure assigning probability one to that point. Since C takes a supremum over all the Dirac measures (and more, since the Dirac measures are a subset of $\mathcal{P}_{practical}$), inserting $\delta(\omega)$ as the probability measure in the supremum expression for C yields $C(\ell) \geq \int_\Omega \ell \; \mathrm{d}\delta(\omega) - 1 \; = \; \sup_{\omega \in \Omega} \ell(\omega) - 1$. Putting these bounds together gives us

$$\sup_{\ell \in B_{simple}, \omega \in \Omega} \ell(\omega) - C(\ell) + C(\ell_0) \;\leq\; \sup_{\ell \in B_{simple}} \sup_{\omega \in \Omega} \ell(\omega) - \left(\sup_{\omega \in \Omega} \ell(\omega) - 1 \right) + \frac{\pi}{4} = 1 + \frac{\pi}{4}.$$

□

This shows that by restricting liabilities to simple functions, this market maker experiences bounded worst-case loss of $1 + \pi/4$. More generally, the cost function in Equation 2 could be scaled by a constant β, yielding a market with a worst-case loss of $\beta(1 + \pi/4)$. Note that setting

$$R(p) = \beta \left(\sum_{\omega \in p_{pp}} p(\omega)^2 - \int_\Omega \arctan\left(\frac{\mathrm{d}p}{\mathrm{d}\lambda}\right) \mathrm{d}\lambda \right),$$

we have that $\sup_{p \in \mathcal{P}_{practical}} R(p) - \inf_{p \in \mathcal{P}_{practical}} R(p) = \beta(1 - (-\arctan(1))) = \beta(1 + \pi/4)$, and so this lost bound matches the bound given in Theorem 2, despite falling outside the general framework.

We think the above discussion offers a path to a greater study of practical cost functions for infinite outcome spaces and reveals some of their unique challenges.

5. COST FUNCTIONS AND SCORING RULES

In this section we relate cost functions for eliciting beliefs to strictly proper scoring rules. Connections between cost functions and scoring rules have been made in prior work for the finite outcome complete market setting [Chen and Pennock 2007; Abernethy et al.

2013] and the incomplete market setting when the outcome and security spaces are both finite [Abernethy and Frongillo 2012]. Our approach generalizes this work and extends it to include markets over infinite outcome spaces by refining the usual notion of strict properness.

A *scoring rule* is a function that maps an expert's forecast and an outcome to a payoff for the expert. For example, if the outcome space is $\Omega = \{\omega_1, \ldots, \omega_n\}$, and an expert's forecast is a probability distribution p over Ω, then when the true outcome $\omega^* \in \Omega$ is revealed, the expert would receive the quantity $s(p, \omega^*)$. Two classic scoring rules for finite outcome spaces are the *log scoring rule* $s(p, \omega_i) = \log(p_i)$ and the *Brier scoring rule* $s(p, \omega_i) = (1 - p_i)^2 + \sum_{j \neq i} p_j^2$. Both satisfy the property that an expert maximizes his expected score by reporting his true beliefs, incentivizing information revelation.

Our discussion of scoring rules will carefully parallel our discussion of cost functions and exploit the same duality. Letting $(\Omega, \mathcal{F})$ be a measurable space and $B \subseteq \mathbf{B}$ a Banach space, we described how market prices correspond to beliefs in P_B. Similarly, we define P_B-scoring rules that accept these "beliefs" as forecasts. Roughly speaking, one can think of reporting an element of P_B as reporting a set of *sufficient statistics* or *expectations of random variables* over Ω. We define a scoring rule

$$s : P_B \times \Omega \to \mathbb{R} \qquad \text{(scoring rule)}$$
$$s_p := s(p_B, \cdot) : \Omega \to \mathbb{R} \ \text{ a function in } B.$$

That is, a P_B-scoring rule is a function mapping a belief from the dual space of B and an outcome to the a score in $\mathbb{R}$, and we require each partial function $s_p = s(p_B, \cdot)$ to be an element of B.[10] Intuitively, we can think of a P_B-scoring rule accepting a prediction $p_B \in P_B$ and returning a *portfolio* $s_p \in B$; when Nature reveals the outcome $\omega^* \in \Omega$, the expert is paid $s_p(\omega^*)$. The *expected score* for reporting $q_B \in P_B$ given beliefs $p \in P$ is then

$$s(q_B, p) := \int_\Omega s_q \mathrm{d}p. \qquad \text{(expected score)}$$

We describe a scoring rule as *proper* if

$$s(p_B, p) \geq s(q_B, p), \quad \forall p \in P, p_B \in P_B \text{ s.t. } p_B \cong_B p, q_B \in P_B \qquad \text{(proper scoring rule)}$$

and *strictly proper* if the inequality is strict whenever $q_B \neq p_B$. If a scoring rule is proper, then an expert maximizes his expected score by reporting his beliefs. If it is strictly proper, he uniquely maximizes his score by doing so.

Theorem 6 gives a characterization of (strictly) proper scoring rules from our duality perspective. This characterization follows that of Gneiting and Raftery [2007], but makes the connection between scoring rules and cost functions explicit. Note that for each $p \in P$, b_{p_B} may be *any* member of the subgradient of R at p_B.

THEOREM 6 ((STRICTLY) PROPER SCORING RULE CHARACTERIZATION). *Let $B \subseteq \mathbf{B}$ be a Banach space and $P \subseteq \mathcal{P}$ a set of probability measures. A P_B-scoring rule $s : P_B \times \Omega \to \mathbb{R}$ is proper if and only if it can be written as*

$$s(p_B, \omega) = b_{p_B}(\omega) - C(b_{p_B})$$

for some function C defined as

$$C(b) = \sup_{p_B \in \mathcal{P}_B} \ \langle b, p_B \rangle - R(p_B)$$

[10] Gneiting and Raftery [2007] considered $P-$scoring rules where P was a convex subset of $\mathcal{P}$ and required each partial function $s(p, \cdot)$ to be $p-$integrable and $P-$quasi-integrable. Letting partial functions be this general requires regularity conditions to be applied to produce a useful characterization and no longer guarantees a scoring rule can be adapted to work in a market.

for some proper, lower semicontinuous, convex function $R : \mathcal{P}_B \to \bar{\mathbb{R}}$ whose subdifferential contains B with $P_B \subseteq \mathrm{dom}(\partial_B R) \subseteq \mathcal{P}_B$, and where $b_{p_B} \in \partial_B R(p_B)$. The scoring rule s is strictly proper if and only if R is also strictly convex where $B-$subdifferentiable.

This theorem shows that the same type of "regularization" function R is used to construct both strictly proper scoring rules and cost functions. There is a slight discrepancy in their duality, however. While a cost function maps liabilities or portfolios to beliefs, a scoring rule maps beliefs to portfolios. The former mapping is many to one, but the latter is one to one. This stems from the fact that a scoring rule is not uniquely identified by the choice of R like a cost function, since the choice of the subgradient associated with each belief also plays a role, whereas in a cost function market, individual traders are free to choose which subgradient (liability function) to associate with their beliefs.

6. CONCLUSION

We have characterized the class of cost function market makers for arbitrary measurable spaces that satisfy a set of intuitive economic properties. This generalizes prior work that considered only finite outcome spaces, and moves beyond what can be achieved in continuous spaces with market makers whose implicit beliefs are restricted to density functions. This characterization demonstrates that we can construct cost functions for continuous random variables with desirable properties like bounded worst-case loss, such as the examples in Section 4. This section also demonstrates a gap between theory and practice, showing that a naive application of our framework can result in odd behavior, and suggests a practical fix. There is considerable room for future work designing practical market makers for intervals and other continuous spaces, and we hope the results here will serve as a starting point.

Fundamental to our analysis is the duality between bounded measurable functions (liabilities or portfolios) and probability measures (beliefs), which arose naturally from the economic properties we required. Our subdifferential analysis explicitly demonstrates the association of a market's liability function and its implicit beliefs.

ACKNOWLEDGMENTS

We are grateful to Miro Dudik for valuable comments on an earlier draft of this paper; the presentation of our main results was greatly improved thanks to his suggestions.

REFERENCES

Abernethy, J., Chen, Y., and Wortman Vaughan, J. 2011. An optimization-based framework for automated market-making. In *ACM EC*.

Abernethy, J., Chen, Y., and Wortman Vaughan, J. 2013. Efficient market making via convex optimization, and a connection to online learning. *ACM Transactions on Economics and Computation*. Forthcoming.

Abernethy, J. and Frongillo, R. 2012. A characterization of scoring rules for linear properties. In *COLT*.

Agrawal, S., Wang, Z., and Ye, Y. 2008. Parimutuel betting on permutations. In *WINE*.

Barbu, V. and Precupanu, T. 2012. *Convexity and Optimization in Banach Spaces*. Springer.

Bauschke, H. H., Borwein, J. M., and Combettes, P. L. 2001. Essential smoothness, essential strict convexity, and Legendre functions in Banach spaces. *Communications in Contemporary Mathematics 3*, 615–647.

Berg, J. E., Forsythe, R., Nelson, F. D., and Rietz, T. A. 2001. Results from a dozen years of election futures markets research. In *Handbook of Experimental Economic Results*, C. A. Plott and V. Smith, Eds.

Brahma, A., Chakraborty, M., Das, S., Lavoie, A., and Magdon-Ismail, M. 2012. A Bayesian market maker. In *ACM EC*.

Chen, Y., Fortnow, L., Nikolova, E., and Pennock, D. M. 2007. Betting on permutations. In *ACM EC*.

Chen, Y. and Pennock, D. M. 2007. A utility framework for bounded-loss market makers. In *UAI*.

Chen, Y., Ruberry, M., and Wortman Vaughan, J. 2012. Designing informative securities. In *UAI*.

Chen, Y. and Wortman Vaughan, J. 2010. A new understanding of prediction markets via no-regret learning. In *ACM EC*.

Das, S. and Magdon-Ismail, M. 2008. Adapting to a market shock: Optimal sequential market-making. In *NIPS*.

Fortnow, L., Kilian, J., Pennock, D. M., and Wellman, M. P. 2004. Betting Boolean-style: A framework for trading in securities based on logical formulas. *Decision Support Systems 39,* 1, 87–104.

Gao, X., Chen, Y., and Pennock, D. M. 2009. Betting on the real line. In *WINE*.

Gao, X. A. and Chen, Y. 2010. An axiomatic characterization of continuous-outcome market makers. In *WINE*.

Ghodsi, M., Mahini, H., Mirrokni, V. S., and Zadimoghaddam, M. 2008. Permutation betting markets: Singleton betting with extra information. In *ACM EC*.

Gneiting, T. and Raftery, A. E. 2007. Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association 102,* 477, 359–378.

Hanson, R. D. 2003. Combinatorial information market design. *Information Systems Frontiers 5,* 1, 107–119.

Hanson, R. D. 2007. Logarithmic market scoring rules for modular combinatorial information aggregation. *Journal of Prediction Markets 1,* 1, 1–15.

Li, X. and Wortman Vaughan, J. 2013. An axiomatic characterization of adaptive-liquidity market makers. In *ACM EC*.

Mangold, B., Dooley, M., Dornfest, R., Flake, G. W., Hoffman, H., Kasturi, T., and Pennock, D. M. 2005. The tech buzz game. *IEEE Computer 38,* 7, 94–97.

Matheson, J. E. and Winkler, R. L. 1976. Scoring rules for continuous probability distributions. *Management Science 22*, 1087–1096.

Minty, G. J. 1964. On the monotonicity of the gradient of a convex function. *Pacific Journal of Mathematics 14*, 243–247.

Ostrovsky, M. 2012. Information aggregation in dynamic markets with strategic traders. *Econometrica 80,* 6, 2595–2648.

Othman, A. and Sandholm, T. 2010. Automated market makers that enable new settings: Extending constant-utility cost functions. In *AMMA*.

Othman, A. and Sandholm, T. 2011. Liquidity-sensitive automated market makers via homogeneous risk measures. In *WINE*.

Othman, A. and Sandholm, T. 2012. Profit-charging market makers with bounded loss, vanishing bid/ask spreads, and unlimited market depth. In *ACM EC*.

Othman, A., Sandholm, T., Pennock, D. M., and Reeves, D. M. 2010. A practical liquidity-sensitive automated market maker. In *ACM EC*.

Pennock, D. M. 2004. A dynamic pari-mutuel market for hedging, wagering, and information aggregation. In *ACM EC*.

Polgreen, P. M., Nelson, F. D., and Neumann, G. R. 2007. Using prediction markets to forecast trends in infectious diseases. *Clinical Infectious Diseases 44,* 2, 272–279.

Ramsey, F. P. 1926. Truth and probability. *The Foundations of Mathematics and Other Logical Essays*, 156–198.

Rockafellar, R. T. 1970. *Convex Analysis*. Princeton University Press.

Wolfers, J. and Zitzewitz, E. 2004. Prediction markets. *Journal of Economic Perspective 18,* 2, 107–126.

A. OMITTED PROOFS

A.1. Proof of Fact 2

This fact appears to be a collection of folk knowledge with precise formal statements and proofs hard to come by in the literature. For example, the relationship between disjoint subdifferentials and strict subgradient inequalities is mentioned by Gneiting and Raftery [2007] and used in their characterization of (strictly) proper scoring rules, yet their citation does not seem to offer a proof of the statement. More recently, an equivalent formal statement of that result appears in Bauschke et al. [2001] although again without an explicit proof. Also, we note that the idea of essential strict convexity [Rockafellar 1970] is similar to our notion of strictly convex where subdifferentiable. In addition to offering a proof, we think our statement usefully collects these equivalences and refines our understanding of the relationship between the subdifferential and strict convexity.

We first prove (1) implies (2). Assume, for a contradiction, that (1) does not imply (2) so there exists $x_0, x_1, x \in \mathrm{dom}(\partial_Y f)$ and $\alpha \in (0,1)$ such that $\alpha x_0 + (1-\alpha)x_1 = x$ and

$$\alpha f(x_0) + (1-\alpha) f(x_1) = f(x)$$

Let $y \in Y$ be the subgradient of f at x. Fact 1 tells us that $f^*(y) = \langle x, y\rangle - f(x)$ and lets us rewrite the above equality as

$$f^*(y) = \alpha\big(\langle x_0, y\rangle - f(x_0)\big) + (1-\alpha)\big(\langle x_1, y\rangle - f(x_1)\big),$$

implying that

$$\langle x, y\rangle - f(x) = \alpha\big(\langle x_0, y\rangle - f(x_0)\big) + (1-\alpha)\big(\langle x_1, y\rangle - f(x_1)\big).$$

Fact 1 also says that x is a solution to the expression $\sup_{x\in X}\ \langle x, y\rangle - f(x)$, which then implies that

$$\langle x_0, y\rangle - f(x_0) = \langle x_1, y\rangle - f(x_1) = \langle x, y\rangle - f(x)$$

and we see that y is a subgradient of f at x_0 and x_1, contradicting our assumption that f has disjoint $Y-$subdifferentials and proving (1) implies (2).

Now we show (2) implies (1) by assuming – again for a contradiction – that (2) does not imply (1). So f is strictly convex where $Y-$subdifferentiable and there exist $x_0, x_1 \in \mathrm{dom}(\partial_Y f)$ such that $\partial_Y f(x_0) \cap \partial_Y f(x_1) \neq \emptyset$. Let y be a common subgradient of x_0 and x_1.

If there exists $x = \alpha x_0 + (1-\alpha)x_1$ for some $\alpha \in (0,1)$ such that the convexity inequality holds with equality then, as in the preceding argument,

$$f^*(y) = \alpha\big(\langle x_0, y\rangle - f(x_0)\big) + (1-\alpha)\big(\langle x_1, y\rangle - f(x_1)\big) = \langle x, y\rangle - f(x)$$

and so Fact 1 tells us this implies y is a subgradient at x, implying f is not strictly convex where $Y-$subdifferentiable—a contradiction.

If there does not exist such an x, then the convexity inequality applies strictly, i.e.,

$$\alpha f(x_0) + (1-\alpha) f(x_1) > f(x),$$

for all $\alpha \in (0,1)$ and all x, x_0, x_1 such that $x = \alpha x_0 + (1-\alpha)x_1$. We can write two subgradient inequalities

$$\begin{aligned} f(x_0) - f(x) &\le \langle x_0 - x, y\rangle && \text{(subgradient inequality at } x_0) \\ f(x_1) - f(x) &\le \langle x_1 - x, y\rangle && \text{(subgradient inequality at } x_1) \end{aligned}$$

and substitute and rearrange them to obtain

$$\begin{aligned} f(x_0) - f(x) &\le \langle x_0 - \alpha x_0 - (1-\alpha)x_1, y\rangle \\ \implies f(x_0) - (1-\alpha)\langle x_0, y\rangle &\le f(x) - (1-\alpha)\langle x_1, y\rangle \end{aligned}$$

and similarly

$$f(x_1) - \alpha\langle x_1, y\rangle \;\leq\; f(x) - \alpha\langle x_0, y\rangle.$$

Adding α times the first expression and $(1-\alpha)$ times the second, we get

$$\begin{aligned} &\alpha f(x_0) - \alpha(1-\alpha)\langle x_0, y\rangle + (1-\alpha)f(x_1) - \alpha(1-\alpha)\langle x_1, y\rangle \\ \leq\;& f(x) - \alpha(1-\alpha)\langle x_1, y\rangle - \alpha(1-\alpha)\langle x_0, y\rangle \end{aligned}$$

and reduce to obtain

$$\alpha f(x_0) + (1-\alpha)f(x_1) \;\leq\; f(x).$$

But we assumed $\alpha f(x_0) + (1-\alpha)f(x_1) > f(x)$ for all $\alpha \in (0,1)$, a contradiction. Thus sharing a subgradient implies the convex inequality holds with the equality at intermediate points, and per the preceding argument this implies a common subgradient in Y. Thus (2) implies (1).

Now we show (3) is equivalent to (1). Let y_0 be a subgradient of f at x_0. If the subgradient inequality is not strict for some x_1 we have

$$\begin{aligned} f(x_0) - f(x_1) &= \langle x_0 - x_1, y_0\rangle \\ \langle x_0, y_0\rangle - f(x_0) &= \langle x_1, y_0\rangle - f(x_1) \end{aligned}$$

And then we apply Fact 1 to see that y_0 is also a subgradient at x_1. Straightforwardly, disjoint $Y-$subdifferentials are equivalent to the subgradient inequality being strict for all $Y-$subgradients.

Finally we note (4) being equivalent to (1) follows immediately from Fact 1 since $f = f^{**}$ with our assumption that f is proper lower semicontinuous and convex. $\square$

A.2. Proof of Lemma 1

We first prove the following auxiliary lemma.

Lemma 2 (CDF Distinguishability). *Any two CDFs F and G on $[0,1]$ such that $\exists x \in [0,1]$ such that $F(x) \neq G(x)$ must differ on a non-empty open set.*

Proof. We begin by showing distinct right-continuous functions differ on a non-empty open set, then applying this results to CDFs.

Let f and g be two right-continuous functions defined on $[a,b) \in \mathbb{R}$. Assume there exists $x \in [a,b)$ such that $f(x) \neq g(x)$. Let $c = f(x) - g(x)$, then by right-continuity there exists $\delta_f, \delta_g > 0$ such that $f(x) - f(x') < c/2$ for all $x' \in (x, x+\delta_f)$, and symmetrically for g. Let $\delta = \min(\delta_f, \delta_g)$, then on the interval $[x, x+\delta)$ f and g are nowhere equal since f is always within $c/2$ of $f(x)$ on that interval and g is always within $c/2$ of $g(x)$, and $f(x)$ and $g(x)$ differ by c, so no number is within $c/2$ of both of them.

Since any two right-continuous functions differ on a non-empty open subset and CDFs are right-continuous if two CDFs F and G differ on $[0,1)$ the result is immediate. If the functions do not differ on $[0,1)$ they do not differ anywhere since the extension of a CDF to $[0,1]$ is unique. $\square$

This lets us take any strictly convex function of the reals and immediately transform it into a strictly convex function of absolutely continuous measures.

Let F and G be the CDFs of two probability measures absolutely continuous with respect to the Lebesgue measure. A Radon-Nikodym derivative (density function) of the measure $\alpha F + (1-\alpha)G$ is then $\alpha\frac{\mathrm{d}F}{\mathrm{d}\lambda} + (1-\alpha)\frac{\mathrm{d}F}{\mathrm{d}\lambda}$. Using the strict convexity of ψ, we have the inequality

$$\psi\left(\alpha\frac{\mathrm{d}F}{\mathrm{d}\lambda} + (1-\alpha)\frac{\mathrm{d}F}{\mathrm{d}\lambda}\right) \;<\; \alpha\psi(\frac{\mathrm{d}F}{\mathrm{d}\lambda}) + (1-\alpha)\psi(\frac{\mathrm{d}F}{\mathrm{d}\lambda})$$

And the same inequality holds for the integrals

$$\int_0^1 \psi\left(\alpha\frac{\mathrm{d}F}{\mathrm{d}\lambda} + (1-\alpha)\frac{\mathrm{d}F}{\mathrm{d}\lambda}\right)\,\mathrm{d}x \;<\; \int_0^1 \alpha\psi\left(\frac{\mathrm{d}F}{\mathrm{d}\lambda}\right) + (1-\alpha)\psi\left(\frac{\mathrm{d}F}{\mathrm{d}\lambda}\right)\,\mathrm{d}x$$

since it holds pointwise and applying Corollary 2 we have that the CDFs differ on an open set and this implies their densities do, too, so the inequality is strict.

Finally, we note that any other Radon-Nikodym derivative differs from the one we constructed only on a Lebesgue-negligible set so the value of any such integral is equivalent and the choice of density function is immaterial to the inequality. □

Proof of Theorem 4

Throughout this proof, let ℓ be a bounded measurable function and Ω_0 and $\Omega_1 \in \mathcal{B}$ two measurable sets such that $\ell(\omega_0) = \ell(\omega_1)$ for all $\omega_0 \in \Omega_0$ and $\omega_1 \in \Omega_1$. Let $p \in \mathcal{P}$ be where the cost function's supremum is attained.

First, assume $\lambda(\Omega_0) = \lambda(\Omega_1) = 0$ and $p(\Omega_0) \neq p(\Omega_1)$. The distribution of p between the two sets does not affect the integral $\int_\Omega \ell\ \mathrm{d}p$, so we only need focus on minimizing $p(\Omega_0)^2 + p(\Omega_1)^2$. Immediately, however, we see this expression admits a minima when the probabilities are equal, an improvement over the assumed supremum and thus a contradiction.

Alternatively, let both $\lambda(\Omega_0)$ and $\lambda(\Omega_1)$ be greater than zero, and p be such that $p(\Omega_0)/\lambda(\Omega_0) \neq p(\Omega_1)/\lambda(\Omega_1)$. Again the integral $\int_\Omega \ell\ \mathrm{d}p$ is unaffected by our choice.

We begin this portion of the proof by demonstrating that since the function ℓ is identical everywhere on Ω_0 and Ω_1, the optimal solution p admits a Radon-Nikodym derivative over both sets. We quickly see that p admits a Radon-Nikodym derivative since having a point mass is simply a penalty, and we assumed p contained only pure point and absolutely continuous parts. This implies we are only interested in minimizing $\int_{\Omega_0+\Omega_1} -\arctan(\mathrm{d}p/\mathrm{d}\lambda)\ \mathrm{d}\lambda$.

Next we show this derivative is uniform. Consider any two points at which to evaluate the derivative, $\mathrm{d}p/\mathrm{d}\lambda(\omega_0)$ and $\mathrm{d}p/\mathrm{d}\lambda(\omega_1)$. Since the $-\arctan$ function is strictly convex, these sum of the $-\arctan$ function evaluated at these two points is minimized when the derivatives are equal. Since the $-\arctan$ function is continuous, this also implies the integral is minimized when all points are the same, and so we see the Radon-Nikodym derivative is uniform across both sets, as desired. Thus we conclude C is Lebesgue unbiased. □

Proof of Theorem 6

Throughout the proof, we use q and p for probability measures in P, q_B and p_B for the elements of the dual space that agree with them, and b_q and b_p for (arbitrary, fixed) elements of the $B-$subdifferential of R at q_B and p_B. We start by showing a scoring rule defined as in the theorem statement is a (strictly) proper P_B-scoring rule.

Letting R be as in the statement. The scoring rule's expected score function is

$$s(q_B, p) = \int_\Omega b_q\ \mathrm{d}p - \sup_{\{\mu | \mu_B \in \mathcal{P}_B\}} \{\int_\Omega b_q\ \mathrm{d}\mu - R(\mu_B)\} = \langle b_q, p\rangle - \langle b_q, q\rangle + R(q_B),$$

for some $b_q \in \partial_B R(q_B)$, where the last equality follows from Fact 1. A scoring rule is proper if

$$\begin{aligned} s(p_B, p) &\geq s(q_B, p),\ \forall q, p \in P \\ \iff R(q_B) - R(p_B) &\leq \langle b_q, q - p\rangle,\ \forall q, p \in P, \end{aligned}$$

which is equivalent to b_q being a subgradient of R at q_B for all q as we assumed; thus s is proper.

Now let R be strictly convex where $B-$subdifferentiable. A scoring rule is strictly proper if the subgradient equation above holds with strict inequality. Fact 2 says this property is equivalent to a function's being strictly convex where $B-$subdifferentiable. Thus we conclude that identifying a scoring rule with such a convex function R is sufficient for it to be (strictly) proper.

Now we show that for every proper P_B-scoring rule there exists a convex function R satisfying the statement's criteria. Since we assumed s is proper,

$$p_B \in \arg\max_{q_B \in P_B} \; s(q_B, p) = \arg\max_{q_B \in P_B} \; \int_\Omega s_q \; \mathrm{d}p, \;\forall p \in P.$$

The partial functions $s(q_B, \cdot)$ are equivalent to bounded measurable functions and so are continuous linear functions of probability measures. We can define $R(p_B) = \sup_{q_B \in P_B} \; s(q_B, p)$ as the pointwise supremum of these functions, which implies it is proper lower semicontinuous and convex. Further, the function s_p is a subgradient of R at p_B since it is a bounded measurable function (by definition) and satisfies the subgradient inequality since

$$R(p_B) - R(q_B) \;\leq\; \langle s_p, p - q\rangle \qquad \text{(subgradient inequality)}$$
$$\iff \int_\Omega s_p \; \mathrm{d}p - \int_\Omega s_q \; \mathrm{d}q \;\leq\; \int_\Omega s_p \; \mathrm{d}(p-q) \iff \int_\Omega s_q \; \mathrm{d}q \;\geq\; \int_\Omega s_p \; \mathrm{d}q$$

and the last inequality holds since s is proper. Further, this subgradient satisfies the theorem statement, since by Fact 1, $C(s_p) = \langle s_p, p\rangle - R(p_B) = 0$. and so $s_p(\omega) - C(s_p) = s_p(\omega)$, and by assumption this subgradient belongs to B.

We conclude by noting that if s is strictly proper the subgradient inequality above is strict and that implies that R is strictly convex where $B-$subdifferentiable. Thus the existence of such a function R is both sufficient and necessary. $\square$