Market Making with Decreasing Utility for Information

Miroslav Dudík
Microsoft Research

Rafael Frongillo
Microsoft Research

Jennifer Wortman Vaughan
Microsoft Research

Abstract

We study information elicitation in cost-function-based combinatorial prediction markets when the market maker's utility for information decreases over time. In the *sudden revelation* setting, it is known that some piece of information will be revealed to traders, and the market maker wishes to prevent guaranteed profits for trading on the sure information. In the *gradual decrease* setting, the market maker's utility for (partial) information decreases continuously over time. We design adaptive cost functions for both settings which: (1) preserve the information previously gathered in the market; (2) eliminate (or diminish) rewards to traders for the publicly revealed information; (3) leave the reward structure unaffected for other information; and (4) maintain the market maker's worst-case loss. Our constructions utilize mixed Bregman divergence, which matches our notion of utility for information.

1 INTRODUCTION

Prediction markets have been used to elicit information in a variety of domains, including business [6, 7, 12, 28], politics [4, 29], and entertainment [25]. In a prediction market, traders buy and sell *securities* with values that depend on some unknown future outcome. For example, a market might offer securities worth \$1 if Norway wins a gold medal in Men's Moguls in the 2014 Winter Olympics and \$0 otherwise. Traders are given an incentive to reveal their beliefs about the outcome by buying and selling securities, e.g., if the current price of the above security is \$0.15, traders who believe that the probability of Norway winning is more than 15% are incentivized to buy and those who believe that the probability is less than 15% are incentivized to sell. The equilibrium price reflects the market consensus about the security's expected payout (which here coincides with the probability of Norway winning the medal).

There has recently been a surge of research on the design of prediction markets operated by a centralized authority called a *market maker*, an algorithmic agent that offers to buy or sell securities at some current price that depends on the history of trades in the market. Traders in these markets can express their belief whenever it differs from the current price by either buying or selling, regardless of whether other traders are willing to act as a counterparty, because the market maker always acts as a counterparty, thus "providing the liquidity" and subsidizing the information collection. This is useful in situations when the lack of interested traders would negatively impact the efficiency in a traditional exchange. Of particular interest to us are *combinatorial prediction markets* [8–10, 17–19, 26] which offer securities on various related events such as "Norway wins a total of 4 gold medals in the 2014 Winter Olympics" and "Norway wins a gold medal in Men's Moguls." In combinatorial markets with large, expressive security spaces, such as an Olympics market with securities covering 88 nations participating in 98 events, the lack of an interested counterparty is a major concern. Only a single trader may be interested in trading the security associated with a specific event, but we would still like the market to incorporate this trader's information.

Most market makers considered in the literature are implemented using a pricing function called the *cost function* [11]. While such markets have many favorable properties [1, 2], the current approaches have several drawbacks that limit their applicability in real-world settings. First, existing work implicitly assumes that the outcome is revealed all at once. When concerned about "just-in-time arbitrage," in which traders closer to the information source make last-minute guaranteed profits by trading on the sure information before the market maker can adjust prices, the market maker can prevent such profits by closing the entire market just before the outcome is revealed. This approach is undesirable when partial information about the outcome is revealed over time, as is often the case in practice, including the Olympics market. For instance, we may learn the results of Men's Moguls before Ladies' Figure Skating has taken place. Closing a large combinatorial market whenever a small portion of the outcome is determined seems to be an unreasonably large intervention.

Second, in real markets, the information captured by the market's consensus prices often becomes less useful as the revelation of the outcome approaches. Consider a market over the event "Unemployment in the U.S. falls below 5.8% by the end of 2015." Although there may be a particular moment when the unemployment rate is publicly revealed, this information becomes gradually less useful as that moment approaches; the government may be less able to act on the information as the end of the year draws near. In the Olympics market, the outcome of a particular competition is often more certain as the final announcement approaches, e.g., if one team is far ahead by the half-time of a hockey game, market forecasts become less interesting. Existing market makers fail to take this diminishing utility for information into account, with the strength of the market incentives remaining constant over time.

To address these two shortcomings of existing markets, we consider two settings:

- a *sudden revelation* setting in which it is known that some piece of information (such as the winner of Men's Moguls) will be publicly revealed at a particular time, driving the market maker's utility for this information to zero; crucially, in this setting we assume that the market maker *does not* have direct access to this information at the time it is revealed, which is realistic in the case of the Olympics where a human might not be available to input winners for all 98 events in real time;
- a *gradual decrease* setting in which the market maker has a diminishing utility for a piece of information (such as the unemployment rate for 2015) over time and therefore is increasingly unwilling to pay for this information even while other information remains valuable.

The sudden revelation setting can be viewed as a special case of the gradual decrease setting. In both cases, we model the relevant information as a variable X, representing a partly determined outcome such as the identity of the gold medal winner in a single sports event.

We consider cost-function-based market makers in which the cost function switches one or many times, and aim to design switching strategies such that: (1) information previously gathered in the market is not lost at the time of the switch, (2) a trader who knows the value of X but has no additional information is unable to profit after the switch (for the sudden revelation setting) or is able to profit less and less over time (in the gradual decrease setting), and (3) the market maker maintains the same reward structure for any other information that traders may have. To formalize these objectives, we define the notion of the market maker's utility (Sec. 2) and show how it corresponds to the *mixed Bregman divergence* [13, 15] (Sec. 2.5).

For the sudden revelation setting (Sec. 3), we introduce a generic cost function switching technique which in many cases removes the rewards for "just-in-time arbitragers" who know only the value of X, while allowing traders with other information to profit, satisfying our objectives.

For the gradual decrease setting (Sec. 4), we focus on *linearly constrained market makers* (LCMMs) [13], proposing a time-sensitive market maker that gradually decreases liquidity by employing the cost function of a different LCMM at each point in time, again meeting our objectives.

Others have considered the design of cost-function-based markets with adaptive liquidity [3, 21–24]. That line of research has typically focused on the goal of slowing down price movement as more money enters the market. In contrast, we adjust liquidity to reflect the current market maker's utility which can be viewed as something external to trading in the market. Additionally, we change liquidity only in the "low-utility" parts of the market, whereas previous work considered market-wide liquidity shifts. Brahma et al. [5] designed a Bayesian market maker that adapts to perceived increases in available information. Our market maker does not try to infer high information periods, but assumes that a schedule of public revelations is given a priori. Our market makers have guaranteed bounds on worst-case loss whereas those of Brahma et al. [5] do not.

2 SETTING AND DESIDERATA

We begin by reviewing cost-function-based market making before describing our desiderata. Here and throughout the paper we make use of many standard results from convex analysis, summarized in Appendix A. All of the proofs in this paper are relegated to the appendix. [1]

2.1 COST-FUNCTION-BASED MARKET MAKING

Let Ω denote the *outcome space*, a finite set of mutually exclusive and exhaustive states of the world. We are interested in the design of cost-function-based market makers operating over a set of K *securities* on Ω specified by a *payoff function* $\boldsymbol{\rho} : \Omega \to \mathbb{R}^K$, where $\boldsymbol{\rho}(\omega)$ denotes the vector of security payoffs if the outcome $\omega \in \Omega$ occurs. Traders may purchase *bundles* $\boldsymbol{r} \in \mathbb{R}^K$ of securities from the market maker, with r_i denoting the quantity of security i that the trader would like to purchase; negative values of r_i are permitted and represent short selling. A trader who purchases a bundle $\boldsymbol{r}$ of securities pays a specified cost for this bundle up front and receives a (possibly negative) payoff of $\boldsymbol{\rho}(\omega) \cdot \boldsymbol{r}$ if the outcome $\omega \in \Omega$ occurs.

Following Chen and Pennock [11] and Abernethy et al. [1, 2], we assume that the market maker initially prices securities using a convex potential function $C : \mathbb{R}^K \to \mathbb{R}$,

[1] The full version of this paper on arXiv includes the appendix.

called the *cost function*. The current state of the market is summarized by a vector $\boldsymbol{q} \in \mathbb{R}^K$, where q_i denotes the total number of shares of security i that have been bought or sold so far. If the market state is $\boldsymbol{q}$ and a trader purchases the bundle $\boldsymbol{r}$, he must pay the market maker $C(\boldsymbol{q}+\boldsymbol{r}) - C(\boldsymbol{q})$. The new market state is then $\boldsymbol{q}+\boldsymbol{r}$. The *instantaneous price* of security i is $\partial C(\boldsymbol{q})/\partial q_i$ whenever well-defined; this is the price per share of an infinitesimally small quantity of security i, and is frequently interpreted as the traders' collective belief about the expected payoff of this security. Any expected payoff must lie in the convex hull of the set $\{\boldsymbol{\rho}(\omega)\}_{\omega\in\Omega}$, called *price space*, denoted $\mathcal{M}$.

While our cost function might not be differentiable at all states $\boldsymbol{q}$, it is always *subdifferentiable* thanks to convexity, i.e., its subdifferential $\partial C(\boldsymbol{q})$ is non-empty for each $\boldsymbol{q}$ and, if it is a singleton, it coincides with the gradient. Let $\boldsymbol{p}(\boldsymbol{q}) \coloneqq \partial C(\boldsymbol{q})$ be called the *price map*. The set $\boldsymbol{p}(\boldsymbol{q})$ is always convex and can be viewed as a multi-dimensional version of the "bid-ask spread". In a state $\boldsymbol{q}$, a trader can make an expected profit if and only if he believes that $\mathbb{E}[\boldsymbol{\rho}(\omega)] \notin \boldsymbol{p}(\boldsymbol{q})$. If C is differentiable at $\boldsymbol{q}$, we slightly abuse notation and also use $\boldsymbol{p}(\boldsymbol{q}) \coloneqq \nabla C(\boldsymbol{q})$.

We assume that the cost function satisfies two standard properties: *no arbitrage* and *bounded loss*. The former means that as long as all outcomes ω are possible, there are no market transactions with a guaranteed profit for a trader. The latter means that the worst-case loss of the market maker is a priori bounded by a constant. Together, they imply that the cost function C can be written in the form $C(\boldsymbol{q}) = \sup_{\boldsymbol{\mu}\in\mathcal{M}}[\boldsymbol{\mu}\cdot\boldsymbol{q} - R(\boldsymbol{\mu})]$, where R is the convex conjugate of C, with $\operatorname{dom} R = \mathcal{M}$. See Abernethy et al. [1, 2] for an analysis of the properties of such markets.

Example 1. *Logarithmic market-scoring rule (LMSR).* The LMSR of Hanson [18, 19] is a cost function for a *complete market* where traders can express any probability distribution over Ω. Here, for any $K \ge 1$, $\Omega = [K] \coloneqq \{1, \dots, K\}$ and $\rho_i(\omega) = \mathbf{1}[i = \omega]$ where $\mathbf{1}[\cdot]$ is a 0/1 indicator, i.e., the security i pays out \$1 if the outcome i occurs and \$0 otherwise. The price space $\mathcal{M}$ is the simplex of probability distributions in K dimensions. The cost function is $C(\boldsymbol{q}) = \ln\bigl(\sum_{i=1}^K e^{q_i}\bigr)$, which is differentiable and generates prices $p_i(\boldsymbol{q}) = e^{q_i}/\bigl(\sum_{j=1}^K e^{q_j}\bigr)$. Here R is the negative entropy function, $R(\boldsymbol{\mu}) = \sum_{i=1}^K \mu_i \ln \mu_i$.

Example 2. *Square.* The square market consists of two independent securities ($K = 2$) each paying out either \$0 or \$1. This can be encoded as $\Omega = \{0,1\}^2$ with $\rho_i(\boldsymbol{\omega}) = \omega_i$ for $i = 1, 2$. The price space is the unit square $\mathcal{M} = [0,1]^2$. Consider the cost function $C(\boldsymbol{q}) = \ln\bigl(1 + e^{q_1}\bigr) + \ln\bigl(1 + e^{q_2}\bigr)$, which is differentiable and generates prices $p_i(\boldsymbol{q}) = e^{q_i}/(1 + e^{q_i})$ for $i = 1, 2$. Using this cost function is equivalent to running two independent binary markets, each with an LMSR cost function. We have $R(\boldsymbol{\mu}) = \sum_{i=1}^2 \mu_i \ln \mu_i + (1 - \mu_i)\ln(1 - \mu_i)$.

PROTOCOL 1: Sudden Revelation Market Makers

Input: initial cost function C, initial state $\boldsymbol{s}^{\mathtt{ini}}$, switch time t, update functions $\mathtt{NewCost}(\boldsymbol{q})$, $\mathtt{NewState}(\boldsymbol{q})$

Until time t:
  sell bundles $\boldsymbol{r}^1, \dots, \boldsymbol{r}^N$ priced using C
    for the total cost $C(\boldsymbol{s}^{\mathtt{ini}}+\boldsymbol{r}) - C(\boldsymbol{s}^{\mathtt{ini}})$ where $\boldsymbol{r} = \sum_{i=1}^N \boldsymbol{r}^i$
  let $\boldsymbol{s} = \boldsymbol{s}^{\mathtt{ini}} + \boldsymbol{r}$
At time t:
  $\tilde{C} \leftarrow \mathtt{NewCost}(\boldsymbol{s})$
  $\tilde{\boldsymbol{s}} \leftarrow \mathtt{NewState}(\boldsymbol{s})$
After time t:
  sell bundles $\tilde{\boldsymbol{r}}^1, \dots, \tilde{\boldsymbol{r}}^{\tilde{N}}$ priced using $\tilde{C}$
    for the total cost $\tilde{C}(\tilde{\boldsymbol{s}} + \tilde{\boldsymbol{r}}) - \tilde{C}(\tilde{\boldsymbol{s}})$ where $\tilde{\boldsymbol{r}} = \sum_{i=1}^{\tilde{N}} \tilde{\boldsymbol{r}}^i$
  let $\tilde{\boldsymbol{s}}^{\mathtt{fin}} = \tilde{\boldsymbol{s}} + \tilde{\boldsymbol{r}}$
Observe ω
Pay $(\boldsymbol{r} + \tilde{\boldsymbol{r}}) \cdot \boldsymbol{\rho}(\omega)$ to traders

PROTOCOL 2: Gradual Decrease Market Makers

Input: time-sensitive cost function $\mathbf{C}(\boldsymbol{q};\, t)$, initial state $\boldsymbol{s}^0$, initial time t^0, update function $\mathtt{NewState}(\boldsymbol{q};\, t, t')$

For $i = 1, \dots, N$ (where N is an unknown number of trades):
  at time $t^i \ge t^{i-1}$: receive a request for a bundle $\boldsymbol{r}^i$
  $\tilde{\boldsymbol{s}}^{i-1} \leftarrow \mathtt{NewState}(\boldsymbol{s}^{i-1};\, t^{i-1}, t^i)$
  sell the bundle $\boldsymbol{r}^i$
    for the cost $\mathbf{C}(\tilde{\boldsymbol{s}}^{i-1} + \boldsymbol{r}^i;\, t^i) - \mathbf{C}(\tilde{\boldsymbol{s}}^{i-1};\, t^i)$
  $\boldsymbol{s}^i \leftarrow \tilde{\boldsymbol{s}}^{i-1} + \boldsymbol{r}^i$
Observe ω
Pay $\sum_{i=1}^N \boldsymbol{r}^i \cdot \boldsymbol{\rho}(\omega)$ to traders

Example 3. *Piecewise linear cost.* Here we describe a non-differentiable cost function for a single binary security ($K = 1$). Let $\Omega = \{0, 1\}$ and $\rho(\omega) = \omega$, so $\mathcal{M} = [0, 1]$. The cost function is $C(q) = \max\{0, q\}$. It gives rise to the price map such that $p(q) = 0$ if $q < 0$, and $p(q) = 1$ if $q > 0$, but at $q = 0$, we have $p(q) = [0, 1]$, i.e., because of non-differentiability we have a bid-ask spread at $q = 0$. Here, $R(\mu) = \mathbb{I}\bigl[\mu \in [0, 1]\bigr]$ where $\mathbb{I}[\cdot]$ is a $0/\infty$ indicator, equal to 0 if true and ∞ if false. This market is uninteresting on its own, but will be useful to us in Sec. 3.3.

2.2 OBSERVATIONS AND ADAPTIVE COSTS

We study two settings. In the *sudden revelation setting*, it is known to both the market maker and the traders that at a particular point in time (the observation time) some information about the outcome (an observation) will be publicly revealed to the traders, but not to the market maker. More precisely, let any function on Ω be called a *random variable* and its value called the *realization* of this random variable. Given a random variable $X : \Omega \to \mathcal{X}$, we assume that its realization is revealed to the traders at the observation time. For a random variable X and a possible realization x, we define the *conditional outcome space* by $\Omega^x \coloneqq \{\omega \in \Omega :\ X(\omega) = x\}$. After observing $X = x$ (where, using standard random variable shorthand, we write X for $X(\omega)$), the traders can conclude that

$\omega \in \Omega^x$. Note that the sets $\{\Omega^x\}_{x\in\mathcal{X}}$ form a partition of Ω.

We design *sudden revelation market makers* (Protocol 1) that replace the cost function C with a new cost function $\tilde{C}$, and the current market state $\boldsymbol{s}$ (i.e., the current value of $\boldsymbol{q}$ in the definition above) with a new market state $\tilde{\boldsymbol{s}}$ in order to reflect the decrease in the utility for information about X. Such a switch would typically occur just before the observation time. Note that we allow the new cost function $\tilde{C}$ as well as the new state $\tilde{\boldsymbol{s}}$ to be chosen adaptively according to the last state $\boldsymbol{s}$ of the original cost function C.

In the *gradual decrease* setting, the utility for information about a future observation X is decreasing continuously over time. We use a *gradual decrease market maker* (Protocol 2) with a time-sensitive cost function $\mathbf{C}(\boldsymbol{q};\, t)$ which sells a bundle $\boldsymbol{r}$ for the cost $\mathbf{C}(\boldsymbol{q}+\boldsymbol{r};\, t)-\mathbf{C}(\boldsymbol{q};\, t)$ at time t, when the market is in a state $\boldsymbol{q}$. We place no assumptions on $\mathbf{C}$ other than that for each t, the function $\mathbf{C}(\cdot;\, t)$ should be an arbitrage-free bounded-loss cost function. The market maker may modify the state between the trades.

Protocol 2 alternates between trades and cost-function switches akin to those in Protocol 1. In each iteration i, the cost function $\mathbf{C}(\cdot;\, t^{i-1})$ is replaced by the cost function $\mathbf{C}(\cdot;\, t^{i})$ while simultaneously replacing the state $\boldsymbol{s}^{i-1}$ by the state $\tilde{\boldsymbol{s}}^{i-1}$. Crucially, unlike Protocol 1, the cost-function switch here is *state independent*, so any state-dependent adaptation happens through the state update. [2]

At a high level, within each of the protocols, our goal is to design switch strategies that satisfy the following criteria:

- Any information that has already been gathered from traders about the relative likelihood of the outcomes in the conditional outcome spaces is preserved.
- A trader who has information about the observation X but has no additional information about the relative likelihood of outcomes in the conditional outcome spaces is unable to profit from this information (for sudden revelation), or the profits of such a trader are decreasing over time (for gradual decrease).
- The market maker continues to reward traders for new information about the relative likelihood of outcomes in the conditional outcome spaces as it did before, with prices reflecting the market maker's utility for information within these sets of outcomes.

To reason about these goals, it is necessary to define what we mean by the information that has been gathered in the market and the market maker's utility.

2.3 MARKET MAKER'S UTILITY

By choosing a cost function, the market maker creates an incentive structure for the traders. Ideally, this incentive structure should be aligned with the market maker's subjective utility for information. That is, the amount the market maker is willing to pay out to traders should reflect the market maker's utility for the information that the traders have provided. In this section, we study how the traders are rewarded for various kinds of information, and use the magnitude of their profits to define the market maker's implicit "utility for information" formally.

We start by defining the market maker's utility for a belief, where a *belief* $\boldsymbol{\mu} \in \mathcal{M}$ is a vector of expected security payoffs $\mathbb{E}[\boldsymbol{\rho}(\omega)]$ for some distribution over Ω.

Definition 1. *The market maker's* utility for a belief $\boldsymbol{\mu} \in \mathcal{M}$ *relative to the state* $\boldsymbol{q}$ *is the maximum expected payoff achievable by a trader with belief* $\boldsymbol{\mu}$ *when the current market state is* $\boldsymbol{q}$*:*

$$\mathtt{Util}(\boldsymbol{\mu};\boldsymbol{q}) \coloneqq \sup\nolimits_{\boldsymbol{r}\in\mathbb{R}^K}\left[\boldsymbol{\mu}\cdot\boldsymbol{r} - C(\boldsymbol{q}+\boldsymbol{r}) + C(\boldsymbol{q})\right] \enspace .$$

Any subset $\mathcal{E} \subseteq \Omega$ is referred to as an *event*. Observations $X = x$ correspond to events Ω^x. Suppose that a trader has observed an event, i.e., a trader knows that $\omega \in \mathcal{E}$, but is otherwise uninformed. The market maker's utility for that event can then be naturally defined as follows.

Definition 2. *The* utility for a (non-null) event $\mathcal{E} \subseteq \Omega$ *relative to the market state* $\boldsymbol{q}$ *is the largest guaranteed payoff that a trader who knows* $\omega \in \mathcal{E}$ *(and has only this information) can achieve when the current market state is* $\boldsymbol{q}$*:*

$$\mathtt{Util}(\mathcal{E};\boldsymbol{q}) \coloneqq \sup_{\boldsymbol{r}\in\mathbb{R}^K}\, \min_{\omega\in\mathcal{E}} \left[\boldsymbol{\rho}(\omega)\cdot\boldsymbol{r} - C(\boldsymbol{q}+\boldsymbol{r}) + C(\boldsymbol{q})\right] \enspace .$$

Finally, consider the setting in which a trader has observed an event $\mathcal{E}$, and also holds a belief $\boldsymbol{\mu}$ consistent with $\mathcal{E}$. Specifically, let $\mathcal{M}(\mathcal{E})$ denote the convex hull of $\{\boldsymbol{\rho}(\omega)\}_{\omega\in\mathcal{E}}$, which is the set of beliefs consistent with the event $\mathcal{E}$, and assume $\boldsymbol{\mu} \in \mathcal{M}(\mathcal{E})$. Then we can define the "excess utility for the belief $\boldsymbol{\mu}$" as the excess utility provided by $\boldsymbol{\mu}$ over just the knowledge of $\mathcal{E}$.

Definition 3. *Given an event* $\mathcal{E}$ *and a belief* $\boldsymbol{\mu} \in \mathcal{M}(\mathcal{E})$*, the* excess utility of $\boldsymbol{\mu}$ over $\mathcal{E}$, *relative to the state* $\boldsymbol{q}$ *is:*

$$\mathtt{Util}(\boldsymbol{\mu} \mid \mathcal{E};\boldsymbol{q}) = \mathtt{Util}(\boldsymbol{\mu};\boldsymbol{q}) - \mathtt{Util}(\mathcal{E};\boldsymbol{q}) \enspace .$$

Note that in these definitions a trader can always choose not to trade ($\boldsymbol{r} = \mathbf{0}$), so the utility for a belief and an event is non-negative. Also it is not too difficult to see that $\mathtt{Util}(\boldsymbol{\mu};\boldsymbol{q}) \geq \mathtt{Util}(\mathcal{E};\boldsymbol{q})$ for any $\boldsymbol{\mu} \in \mathcal{M}(\mathcal{E})$, so the excess utility for a belief is also non-negative.

In Sec. 2.5, we show that given a state $\boldsymbol{q}$ and a non-null event $\mathcal{E}$, there always exists a (possibly non-unique) belief $\boldsymbol{\mu} \in \mathcal{E}$ such that $\mathtt{Util}(\boldsymbol{\mu} \mid \mathcal{E};\boldsymbol{q}) = 0$. Thus, a trader with such a "worst-case" belief is able to achieve in expectation no reward beyond what any trader that just observed $\mathcal{E}$ would receive. We show that these worst-case beliefs correspond to certain kinds of "projections" of the current price $\boldsymbol{p}(\boldsymbol{q})$ onto $\mathcal{M}(\mathcal{E})$. For LMSR, the projections are with respect to KL divergence and correspond to the usual conditional probability distributions. Moreover, for sufficiently

[2] This simplifying restriction matches our solution concept in Sec. 4, but it could be dropped for greater generality.

Table 1: Information Desiderata

PRICE	*Preserve prices:* $\tilde{\boldsymbol{p}}(\tilde{\boldsymbol{s}}) = \boldsymbol{p}(\boldsymbol{s})$.
CONDPRICE	*Preserve conditional prices:* $\tilde{\boldsymbol{p}}(X{=}x;\tilde{\boldsymbol{s}}) = \boldsymbol{p}(X{=}x;\boldsymbol{s}) \quad \forall x \in \mathcal{X}$.
DECUTIL	*Decrease profits for uninformed traders:* $\tilde{\texttt{U}}\texttt{til}(X{=}x;\tilde{\boldsymbol{s}}) \leq \texttt{Util}(X{=}x;\boldsymbol{s}) \quad \forall x \in \mathcal{X}$, with sharp inequality if $\texttt{Util}(X{=}x;\boldsymbol{s}) > 0$.
ZEROUTIL	*No profits for uninformed traders:* $\tilde{\texttt{U}}\texttt{til}(X{=}x;\tilde{\boldsymbol{s}}) = 0 \quad \forall x \in \mathcal{X}$.
EXUTIL	*Preserve excess utility:* $\tilde{\texttt{U}}\texttt{til}(\boldsymbol{\mu}\vert X{=}x;\tilde{\boldsymbol{s}}) = \texttt{Util}(\boldsymbol{\mu}\vert X{=}x;\boldsymbol{s})$ for all $x \in \mathcal{X}$ and $\boldsymbol{\mu} \in \mathcal{M}(X{=}x)$.

smooth cost functions (including LMSR) they correspond to market prices that result when a trader is optimizing his guaranteed profit from the information $\omega \in \mathcal{E}$ as in Definition 2 (see Appendix E). Because of this motivation, such beliefs are referred to as "conditional price vectors."

Definition 4. *A vector $\boldsymbol{\mu} \in \mathcal{M}(\mathcal{E})$ is called a* conditional price vector, *conditioned on $\mathcal{E}$, relative to the state $\boldsymbol{q}$ if* $\texttt{Util}(\boldsymbol{\mu};\boldsymbol{q}) = \texttt{Util}(\mathcal{E};\boldsymbol{q})$. *The set of such conditional price vectors is denoted*

$$\boldsymbol{p}(\mathcal{E};\boldsymbol{q}) \coloneqq \{\boldsymbol{\mu} \in \mathcal{M}(\mathcal{E}) : \texttt{Util}(\boldsymbol{\mu};\boldsymbol{q}) = \texttt{Util}(\mathcal{E};\boldsymbol{q})\} \ .$$

See Appendix F for additional motivation for our definitions of utility and conditioning. With these notions defined, we can now state our desiderata.

2.4 DESIDERATA

Recall that we aim to design mechanisms which replace a cost function C at a state $\boldsymbol{s}$, with a new cost function $\tilde{C}$ at a state $\tilde{\boldsymbol{s}}$. Let $\texttt{Util}$ denote the utility for information with respect to C and $\tilde{\texttt{U}}\texttt{til}$ with respect to $\tilde{C}$, and let $\boldsymbol{p}$ and $\tilde{\boldsymbol{p}}$ be the respective price maps. In our mechanisms, we attempt to satisfy (a subset of) the conditions on information structures as listed in Table 1.

Conditions PRICE and CONDPRICE capture the requirement to preserve the information gathered in the market. The current price $\boldsymbol{p}(\boldsymbol{q})$ is the ultimate information content of the market at a state $\boldsymbol{q}$ *before* the observation time, but it is not necessarily the right notion of information content *after* the observation time. When we do not know the realization x, we may wish to set up the market so that any trader who has observed $X = x$ and would like to maximize the guaranteed profit would move the market to the same conditional price vector as in the previous market. This is captured by CONDPRICE.

DECUTIL models a scenario in which the utility for information about X decreases over time, and ZEROUTIL represents the extreme case in which utility decreases to zero. These conditions are in friction with EXUTIL, which aims to maintain the utility structure over the conditional outcome spaces. A key challenge is to satisfy EXUTIL and ZEROUTIL (or DECUTIL) simultaneously.

Apart from the information desiderata of Table 1, we would like to maintain an important feature of cost-function-based market makers: their ability to bound the worst-case loss to the market maker. Specifically, we would like to show that there is some *finite* bound (possibly depending on the initial state) such that no matter what trades are executed and which outcome ω occurs, the market maker will lose no more than the amount of the bound. It turns out that the solution concepts introduced in this paper maintain the same loss bound as guaranteed for using just the market's original cost function C, but since the focus of the paper is on the information structures, worst-case loss analysis is relegated to Appendix H.

In Sec. 3, we study in detail the sudden revelation setting with the goal of instantiating Protocol 1 in a way that achieves ZEROUTIL while satisfying CONDPRICE and EXUTIL. Our key result is a characterization and a geometric sufficient condition for when this is possible.

In Sec. 4, we examine instantiations of Protocol 2 for the gradual decrease setting. Our construction focuses on linearly-constrained market makers (LCMM) [13], which naturally decompose into submarkets. We show how to achieve PRICE, CONDPRICE, DECUTIL and EXUTIL in LCMMs. We also show that it is possible to simultaneously decrease the utility for information in each submarket according to its own schedule, while maintaining PRICE.

Before we develop these mechanisms, we introduce the machinery of Bregman divergences, which helps us analyze notions of utility for information.

2.5 BREGMAN DIVERGENCE AND UTILITY

To analyze the market maker's utility for information, we show how it corresponds to a specific notion of distance built into the cost function, the *mixed (or generalized) Bregman divergence* [13, 15]. Let R be the conjugate of C.[3] The mixed Bregman divergence between a belief $\boldsymbol{\mu}$ and a state $\boldsymbol{q}$ is defined as $D(\boldsymbol{\mu}\|\boldsymbol{q}) \coloneqq R(\boldsymbol{\mu}) + C(\boldsymbol{q}) - \boldsymbol{q}\cdot\boldsymbol{\mu}$. The conjugacy of R and C implies that $D(\boldsymbol{\mu}\|\boldsymbol{q}) \geq 0$ with equality iff $\boldsymbol{\mu} \in \partial C(\boldsymbol{q}) = \boldsymbol{p}(\boldsymbol{q})$, i.e., if the price vector "matches" the state (see Appendix A). The geometric interpretation of mixed Bregman divergence is as a gap between a tangent and the graph of the function R (see Fig. 1).

To see how the divergence relates to traders' beliefs, consider a trader who believes that $\mathbb{E}[\boldsymbol{\rho}(\omega)] = \boldsymbol{\mu}'$ and moves the market from state $\boldsymbol{q}$ to state $\boldsymbol{q}'$. The expected payoff to this trader is $(\boldsymbol{q}' - \boldsymbol{q})\cdot\boldsymbol{\mu}' - C(\boldsymbol{q}') + C(\boldsymbol{q}) = D(\boldsymbol{\mu}'\|\boldsymbol{q}) - D(\boldsymbol{\mu}'\|\boldsymbol{q}')$. This payoff increases as $D(\boldsymbol{\mu}'\|\boldsymbol{q}')$

[3] The conjugate is also, less commonly, called the "dual".

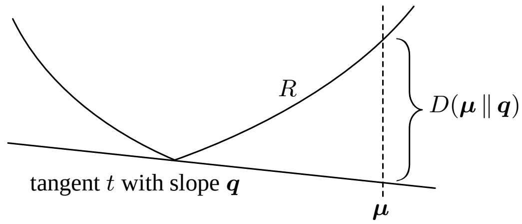

Figure 1: The mixed Bregman divergence $D(\boldsymbol{\mu}\|\boldsymbol{q})$ derived from the conjugate pair C and R measures the distance between the tangent with slope $\boldsymbol{q}$ and the value of R evaluated at $\boldsymbol{\mu}$. By conjugacy, the tangent t is described by $t(\boldsymbol{\mu}) = \boldsymbol{\mu}\cdot\boldsymbol{q} - C(\boldsymbol{q})$. Note that the divergence is well defined even when R is not differentiable, because each slope vector determines a unique tangent.

decreases. Thus, subject to the trader's budget constraints, the trader is incentivized to move to the state $\boldsymbol{q}'$ which is as "close" to his/her belief $\boldsymbol{\mu}'$ as possible in the sense of a smaller value $D(\boldsymbol{\mu}'\|\boldsymbol{q}')$, with the largest expected payoff when $D(\boldsymbol{\mu}'\|\boldsymbol{q}') = 0$. This argument shows that $D(\cdot\|\cdot)$ is an implicit measure of distance used by traders.

The next theorem shows that the Bregman divergence also matches the concepts defined in Sec. 2.3. Specifically, we show that (1) the utility for a belief coincides with the Bregman divergence, (2) the utility for an event $\mathcal{E}$ is the smallest divergence between the current market state and $\mathcal{M}(\mathcal{E})$, and (3) the conditional price vector is the (Bregman) projection of the current market state on $\mathcal{M}(\mathcal{E})$, i.e., it is a belief in $\mathcal{M}(\mathcal{E})$ that is "closest to" the current market state.

Theorem 1. *Let $\boldsymbol{\mu} \in \mathcal{M}$, $\boldsymbol{q} \in \mathbb{R}^K$ and $\emptyset \neq \mathcal{E} \subseteq \Omega$. Then*

$$\texttt{Util}(\boldsymbol{\mu}; \boldsymbol{q}) = D(\boldsymbol{\mu}\|\boldsymbol{q}) \ , \tag{1}$$

$$\texttt{Util}(\mathcal{E}; \boldsymbol{q}) = \min_{\boldsymbol{\mu}'\in\mathcal{M}(\mathcal{E})} D(\boldsymbol{\mu}'\|\boldsymbol{q}) \ , \tag{2}$$

$$\boldsymbol{p}(\mathcal{E}; \boldsymbol{q}) = \operatorname{argmin}_{\boldsymbol{\mu}'\in\mathcal{M}(\mathcal{E})} D(\boldsymbol{\mu}'\|\boldsymbol{q}) \ . \tag{3}$$

We finish this section by characterizing when ExUtil is satisfied and showing that it implies CondPrice. Recall that $\Omega^x = \{\omega : \ X(\omega) = x\}$ and let $\mathcal{M}^x := \mathcal{M}(\Omega^x)$.

Proposition 1. ExUtil *holds if and only if for all $x \in \mathcal{X}$, there exists some c^x such that for all $\boldsymbol{\mu} \in \mathcal{M}^x$, $D(\boldsymbol{\mu}\|\boldsymbol{s}) - \tilde{D}(\boldsymbol{\mu}\|\tilde{\boldsymbol{s}}) = c^x$. Moreover,* ExUtil *implies* CondPrice.

3 SUDDEN REVELATION

In this section, we consider the design of sudden revelation market makers (Protocol 1). In this setting, partial information in the form of the realization of X is revealed to market participants (but not to the market maker) at a predetermined time, as might be the case if the medal winners of an Olympic event are announced but no human is available to input this information into the automated market maker on behalf of the market organizer. The random variable X and the observation time are assumed to be known, and the market maker wishes to "close" the submarket with respect to X just before the observation time, without knowing the realization x, while leaving the rest of the market unchanged.

Stated in terms of our formalism, we wish to find functions `NewState` and `NewCost` from Protocol 1 such that the desiderata CondPrice, ExUtil, and ZeroUtil from Table 1 are satisfied. This implies that traders who know only that $X = x$ are not rewarded after the observation time, but traders with new information about the outcome space conditioned on $X = x$ are rewarded exactly as before. As a result, trading immediately resumes in a "conditional market" on $\mathcal{M}(\Omega^x)$ for the correct realization x, without the market maker needing to know x and without any other human intervention. We refer to the goal of simultaneously achieving CondPrice, ExUtil, and ZeroUtil as achieving *implicit submarket closing*.

For convenience, throughout this section we write $\mathcal{M}^x := \mathcal{M}(\Omega^x)$ to denote the conditional price space, and $\mathcal{M}^\star := \bigcup_{x\in\mathcal{X}} \mathcal{M}^x$ to denote prices possible after the observation.

3.1 SIMPLIFYING THE OBJECTIVE

We first show that achieving implicit submarket closing can be reduced to finding a function $\tilde{R}$ satisfying a simple set of constraints, and defining `NewCost` to return the conjugate $\tilde{C}$ of $\tilde{R}$. As a first step, we observe that it is without loss of generality to let `NewState` be an identity map, i.e., to assume that $\tilde{\boldsymbol{s}} = \boldsymbol{s}$; when this is not the case, we can obtain an equivalent market by setting $\tilde{\boldsymbol{s}} = \boldsymbol{s}$ and shifting $\tilde{C}$ so that the Bregman divergence is unchanged.

Lemma 1. *Any desideratum of Table 1 holds for $\tilde{C}$ and $\tilde{\boldsymbol{s}}$ if and only if it holds for $\tilde{C}'(\boldsymbol{q}) = \tilde{C}(\boldsymbol{q} + \tilde{\boldsymbol{s}} - \boldsymbol{s})$ and $\tilde{\boldsymbol{s}}' = \boldsymbol{s}$.*

To simplify exposition, we assume that $\tilde{\boldsymbol{s}} = \boldsymbol{s}$ throughout the rest of the section as we search for conditions on `NewCost` that achieve implicit submarket closing. Under this assumption, Proposition 1 can be used to characterize our goal in terms of $\tilde{R}$. Specifically, we show that ExUtil and CondPrice hold if $\tilde{R}$ differs from R by a (possibly different) constant on each conditional price space $\mathcal{M}^x$.

Lemma 2. *When $\tilde{\boldsymbol{s}} = \boldsymbol{s}$,* ExUtil *and* CondPrice *hold together if and only if there exist constants b^x for $x \in \mathcal{X}$ such that $\tilde{R}(\boldsymbol{\mu}) = R(\boldsymbol{\mu}) - b^x$ for all $x \in \mathcal{X}$ and $\boldsymbol{\mu} \in \mathcal{M}^x$.*

This suggests parameterizing our search for $\tilde{R}$ by vectors $\boldsymbol{b} = \{b^x\}_{x\in\mathcal{X}}$. For $\boldsymbol{b} \in \mathbb{R}^{\mathcal{X}}$, define a function

$$R^{\boldsymbol{b}}(\boldsymbol{\mu}) = \begin{cases} R(\boldsymbol{\mu}) - b^x & \text{if } \boldsymbol{\mu} \in \mathcal{M}^x, x \in \mathcal{X}, \\ \infty & \text{otherwise.} \end{cases}$$

If the sets $\mathcal{M}^x$ overlap, $R^{\boldsymbol{b}}$ is not well defined for all $\boldsymbol{b}$. Whenever we write $R^{\boldsymbol{b}}$, we assume that $\boldsymbol{b}$ is such that $R^{\boldsymbol{b}}$ is well defined. To satisfy Lemma 2 with a specific $\boldsymbol{b}$, it suffices to find a convex function $\tilde{R}$ "consistent with" $R^{\boldsymbol{b}}$ in the following sense.

Definition 5. *We say that a function $\tilde{R}$ is* consistent *with $R^{\boldsymbol{b}}$ if $\tilde{R}(\boldsymbol{\mu}) = R^{\boldsymbol{b}}(\boldsymbol{\mu})$ for all $\boldsymbol{\mu} \in \mathcal{M}^\star$.*

We next simplify our objective further by proving that whenever implicit submarket closing is achievable, it suffices to consider functions `NewCost` that set $\tilde{C}$ to be the conjugate of the largest convex function consistent with $R^{\boldsymbol{b}}$ for some $\boldsymbol{b} \in \mathbb{R}^{\mathcal{X}}$. To establish this, we examine properties of the *convex roof* of $R^{\boldsymbol{b}}$, the largest convex function that lower-bounds (but is not necessarily consistent with) $R^{\boldsymbol{b}}$.

Definition 6. *Given a function $f : \mathbb{R}^K \to (-\infty, \infty]$, the* convex roof *of f, denoted $(\operatorname{conv} f)$, is the largest convex function lower-bounding f, defined by*

$$(\operatorname{conv} f)(\boldsymbol{x}) := \sup \{g(\boldsymbol{x}) : g \in \mathcal{G}, g \le f\}$$

where $\mathcal{G}$ is the set of convex functions $g : \mathbb{R}^K \to (-\infty, \infty]$, and the condition $g \le f$ holds pointwise.

The convex roof is analogous to a convex hull, and the epigraph of $(\operatorname{conv} f)$ is the convex hull of the epigraph of f. See Hiriart-Urruty and Lemaréchal [30, §B.2.5] for details.

Example 4. Recall the square market of Example 2. Let $X(\boldsymbol{\omega}) = \omega_1$, so traders observe the payoff of the first security at observation time. Then $\mathcal{M}^x = \{x\} \times [0, 1]$ for $x \in \{0, 1\}$. For simplicity, let $\boldsymbol{b} = \boldsymbol{0}$. We have $R^{\boldsymbol{b}}(\boldsymbol{\mu}) = \mu_2 \ln \mu_2 + (1-\mu_2) \ln(1-\mu_2)$ for $\boldsymbol{\mu} \in \mathcal{M}^1 \cup \mathcal{M}^2$ and $R^{\boldsymbol{b}}(\boldsymbol{\mu}) = \infty$ for all other $\boldsymbol{\mu}$. Examining the convex hull of the epigraph of $R^{\boldsymbol{b}}$ gives us that for all $\boldsymbol{\mu} \in [0, 1]^2$, we have $(\operatorname{conv} R^{\boldsymbol{b}})(\boldsymbol{\mu}) = \mu_2 \ln \mu_2 + (1 - \mu_2) \ln(1 - \mu_2)$.

As this example illustrates, the roof of $R^{\boldsymbol{b}}$ is the "flattest" convex function lower-bounding $R^{\boldsymbol{b}}$. Given the geometric interpretation of Bregman divergence (Fig. 1), a "flatter" $\tilde{R}$ yields a smaller utility for information. This flatness plays a key role in achieving ZEROUTIL. Assume that $\tilde{R}$ is consistent with $R^{\boldsymbol{b}}$, so CONDPRICE and EXUTIL hold by Lemma 2. Following the intuition in Fig. 1, to achieve ZEROUTIL, i.e., $\tilde{D}(\hat{\boldsymbol{\mu}}^x \| \boldsymbol{s}) = 0$ across all $x \in \mathcal{X}$ and $\hat{\boldsymbol{\mu}}^x \in \boldsymbol{p}(\Omega^x; \boldsymbol{s})$, it must be the case that for all x and $\hat{\boldsymbol{\mu}}^x$, the function values $\tilde{R}(\hat{\boldsymbol{\mu}}^x)$ lie on the tangent of $\tilde{R}$ with slope $\boldsymbol{s}$. That is, the graph of $\tilde{R}$ needs to be *flat* across the points $\hat{\boldsymbol{\mu}}^x$. This suggests that the roof might be a good candidate for $\tilde{R}$. This intuition is formalized in the following lemma, which states that instead of considering arbitrary convex $\tilde{R}$ consistent with $R^{\boldsymbol{b}}$, we can consider $\tilde{R}$ which take the form of a convex roof.

Lemma 3. *If any convex function $\tilde{R}$ is consistent with $R^{\boldsymbol{b}}$ then so is the convex roof $\tilde{R}' = (\operatorname{conv} R^{\boldsymbol{b}})$. Furthermore, if $\tilde{R}$ satisfies* ZEROUTIL *or* DECUTIL *then so does $\tilde{R}'$.*

3.2 IMPLICIT SUBMARKET CLOSING

We now have the tools to answer the central question of this section: When can we achieve implicit submarket closing? Lemma 1 implies that we can assume that `NewState` is the identity function, and Lemmas 2 and 3 imply that it suffices to consider functions `NewCost` that set $\tilde{C}$ to the conjugate of $\tilde{R} = (\operatorname{conv} R^{\boldsymbol{b}})$ for some $\boldsymbol{b} \in \mathbb{R}^{\mathcal{X}}$. What remains is to find the vector $\boldsymbol{b}$ that guarantees ZEROUTIL. As mentioned above, ZEROUTIL is satisfied if and only if $\left(\hat{\boldsymbol{\mu}}^x, \tilde{R}(\hat{\boldsymbol{\mu}}^x)\right)$ lies on the tangent of $\tilde{R}$ with the slope $\boldsymbol{s}$ for all $x \in \mathcal{X}$ and $\hat{\boldsymbol{\mu}}^x \in \boldsymbol{p}(\Omega^x; \boldsymbol{s})$. This implies that $\tilde{R}(\hat{\boldsymbol{\mu}}^x) = \hat{\boldsymbol{\mu}}^x \cdot \boldsymbol{s} - c$ for all x and $\hat{\boldsymbol{\mu}}^x$ and some constant c. The specific choice of c does not matter since $\tilde{D}$ is unchanged by vertical shifts of the graph of $\tilde{R}$. For convenience, we set $c = C(\boldsymbol{s})$, which makes the tangents of R and $\tilde{R}$ with the slope $\boldsymbol{s}$ coincide. This and Lemma 2 then yield the choice of $\boldsymbol{b} = \hat{\boldsymbol{b}}$, with

$$\hat{b}^x := R(\hat{\boldsymbol{\mu}}^x) + C(\boldsymbol{s}) - \hat{\boldsymbol{\mu}}^x \cdot \boldsymbol{s} = D(\hat{\boldsymbol{\mu}}^x \| \boldsymbol{s}) \qquad (4)$$

for all x and any choice of $\hat{\boldsymbol{\mu}}^x \in \boldsymbol{p}(\Omega^x; \boldsymbol{s})$. The resulting construction of $\tilde{R} = (\operatorname{conv} R^{\hat{\boldsymbol{b}}})$ can be described using geometric intuition. First, consider the tangent of R with slope equal to the current market state $\boldsymbol{s}$. For each $x \in \mathcal{X}$, take the subgraph of R over the set $\mathcal{M}^x$ and let it "fall" vertically until it touches this tangent at the point $\hat{\boldsymbol{\mu}}^x$. The set of fallen graphs for all x together describes $R^{\hat{\boldsymbol{b}}}$ and the convex hull of the fallen epigraphs yields $\tilde{R} = (\operatorname{conv} R^{\hat{\boldsymbol{b}}})$.

Defining `NewCost` using this construction guarantees ZEROUTIL, but CONDPRICE and EXUTIL are achieved only when $\tilde{R}$ is consistent with $R^{\hat{\boldsymbol{b}}}$. Conversely, whenever the three properties are achievable, this construction produces a function $\tilde{R}$ consistent with $R^{\hat{\boldsymbol{b}}}$. This yields a full characterization of when implicit submarket closing is achievable.

Theorem 2. *Let $\hat{\boldsymbol{b}}$ be defined as in Eq. (4).* CONDPRICE, EXUTIL, *and* ZEROUTIL *can be satisfied using Protocol 1 if and only if $(\operatorname{conv} R^{\hat{\boldsymbol{b}}})$ is consistent with $R^{\hat{\boldsymbol{b}}}$. In this case, they can be achieved with `NewState` as the identity and `NewCost` outputting the conjugate of $\tilde{R} = (\operatorname{conv} R^{\hat{\boldsymbol{b}}})$.*

3.3 CONSTRUCTING THE COST FUNCTION

Theorem 2 describes how to achieve implicit submarket closing by defining the cost function $\tilde{C}$ output by `NewCost` implicitly via its conjugate $\tilde{R}$. In this section, we provide an explicit construction of the resulting cost function, and illustrate the construction through examples.

Fixing R, for each $x \in \mathcal{X}$ define a function $C^x(\boldsymbol{q}) := \sup_{\boldsymbol{\mu} \in \mathcal{M}^x} \left[\boldsymbol{q} \cdot \boldsymbol{\mu} - R(\boldsymbol{\mu})\right]$. Each function C^x can be viewed as a bounded-loss and arbitrage-free cost function for outcomes in Ω^x. The conjugate of each C^x coincides with R on $\mathcal{M}^x$ (and is infinite outside $\mathcal{M}^x$). The explicit expression for $\tilde{C}$ is described in the following proposition.

Proposition 2. *For a given C with conjugate R, define $\hat{\boldsymbol{b}}$ as in Eq. (4) and let $\tilde{R} = (\operatorname{conv} R^{\hat{\boldsymbol{b}}})$. The conjugate $\tilde{C}$ of $\tilde{R}$ can be written $\tilde{C}(\boldsymbol{q}) = \max_{x \in \mathcal{X}} \left[\hat{b}^x + C^x(\boldsymbol{q})\right]$. Furthermore, for each $x \in \mathcal{X}$, $\hat{b}^x = C(\boldsymbol{s}) - C^x(\boldsymbol{s})$.*

At any market state $\boldsymbol{q}$ with a unique $\hat{x} := \operatorname{argmax}_{x \in \mathcal{X}} \left[\hat{b}^x + C^x(\boldsymbol{q})\right]$, the price according to $\tilde{C}$ lies in the set $\mathcal{M}^{\hat{x}}$. When $\hat{x}$ is not unique, the market has a bid-ask spread. The addition of $\hat{b}^x$ ensures that the bid-ask spread at the market state $\boldsymbol{s}$ contains conditional prices $\hat{\boldsymbol{\mu}}^x$ across all x. To illustrate this construction, we return to the example of a square.

Example 5. Consider again the square market from Examples 2 and 4 with $X(\boldsymbol{\omega}) = \omega_1$. One can verify that $C^x(\boldsymbol{q}) = xq_1 + \ln\bigl(1 + e^{q_2}\bigr)$ for $x \in \{0, 1\}$. Prop. 2 gives

$$\begin{aligned}\tilde{C}(\boldsymbol{q}) &= \max_{x\in\{0,1\}}\Bigl[x(q_1 - s_1) + \ln(1+e^{q_2}) + \ln(1+e^{s_1})\Bigr] \\ &= \max\{0, q_1 - s_1\} + \ln(1 + e^{s_1}) + \ln(1 + e^{q_2}).\end{aligned}$$

In switching from C to $\tilde{C}$ we have effectively changed the first term of our cost from a basic LMSR cost for a single binary security to the piecewise linear cost of Example 3, introducing a bid-ask spread for security 1 when $q_1 = s_1$; states $\boldsymbol{q} = (s_1, q_2)$ have $\tilde{\boldsymbol{p}}(\boldsymbol{q}) = [0, 1] \times \{e^{q_2}/(1 + e^{q_2})\}$. The market for security 1 has thus implicitly closed; as the new market begins with $\boldsymbol{q} = \boldsymbol{s}$, any trader can switch the price of security 1 to 0 or 1 by simply purchasing an infinitesimal quantity of security 1 in the appropriate direction, at essentially no cost and with no ability to profit.

The example above illustrates our cost function construction, but does not show that $\tilde{R}$ is consistent with $R^{\hat{\boldsymbol{b}}}$ as required by Theorem 2. In fact, it is consistent. This follows from the sufficient condition proved in Appendix G.2. Briefly, the condition is that $\mathcal{M}^\star$ does not contain any price vectors $\boldsymbol{\mu}$ that can be expressed as nontrivial convex combinations of vectors from multiple $\mathcal{M}^x$.

In Appendix G.3, we show that this sufficient condition applies to many settings of interest such as arbitrary partitions of simplex and submarket observations in binary-payoff LCMMs (defined in Sec. 4), which were used to run a combinatorial market for the 2012 U.S. Elections [14].

A case in which the sufficient condition is violated is the square market with $X(\boldsymbol{\omega}) = \omega_1 + \omega_2 \in \{0, 1, 2\}$, where $\mathcal{M}^0 = (0, 0)$ and $\mathcal{M}^2 = (1, 1)$ but $(\frac{1}{2}, \frac{1}{2}) = \frac{1}{2}(0, 0) + \frac{1}{2}(1, 1) \in \mathcal{M}^1$. This particular example also fails to satisfy Theorem 2 (see Appendix G.1), but in general the sufficient condition is not necessary (see Appendix G.4).

4 GRADUAL DECREASE

We now consider gradual decrease market makers (Protocol 2) for the gradual decrease setting in which the utility of information about a future observation X is decreasing continuously over time. We focus on *linearly constrained market makers* (LCMMs) [13], which naturally decompose into submarkets. Our proposed gradual decrease market maker employs a different LCMM at each time step, and satisfies various desiderata of Sec. 2.4 between steps.

As a warm-up for the concepts introduced in this section, we show how the "liquidity parameter" can be used to implement a decreasing utility for information.

Example 6. *Homogeneous decrease in utility for information.* We begin with a differentiable cost function C in a state $\boldsymbol{s}$. Let $\alpha \in (0, 1)$, and define $\tilde{C}(\boldsymbol{q}) = \alpha C(\boldsymbol{q}/\alpha)$, and $\tilde{\boldsymbol{s}} = \alpha \boldsymbol{s}$. $\tilde{C}$ is parameterized by the "liquidity parameter" α. The transformation $\tilde{\boldsymbol{s}}$ guarantees the preservation of prices, i.e., $\tilde{\boldsymbol{p}}(\tilde{\boldsymbol{s}}) = \nabla\tilde{C}(\tilde{\boldsymbol{s}}) = \alpha\nabla C(\tilde{\boldsymbol{s}}/\alpha)/\alpha = \nabla C(\boldsymbol{s}) = \boldsymbol{p}(\boldsymbol{s})$. We can derive that $\tilde{R}(\boldsymbol{\mu}) = \alpha R(\boldsymbol{\mu})$, and $\tilde{D}(\boldsymbol{\mu}\|\boldsymbol{q}) = \alpha D(\boldsymbol{\mu}\|\boldsymbol{q}/\alpha)$, so, for all $\boldsymbol{\mu}$, $\tilde{D}(\boldsymbol{\mu}\|\tilde{\boldsymbol{s}}) = \alpha D(\boldsymbol{\mu}\|\boldsymbol{s})$. In words, the utility for all beliefs $\boldsymbol{\mu}$ with respect to the current state is decreased according to the multiplier α.

This idea will be the basis of our construction. We next define the components of our setup and prove the desiderata.

4.1 LINEARLY CONSTRAINED MARKETS

Recall that $\boldsymbol{\rho} : \Omega \to \mathbb{R}^K$ is the payoff function. Let $\mathcal{G}$ be a system of non-empty disjoint subsets $g \subseteq [K]$ forming a partition of coordinates of $\boldsymbol{\rho}$, so $[K] = \bigcup_{g\in\mathcal{G}} g$. We use the notation $\boldsymbol{\rho}_g(\omega) := (\rho_i(\omega))_{i\in g}$ for the block of coordinates in g, and similarly $\boldsymbol{\mu}_g$ and $\boldsymbol{q}_g$. Blocks g describe groups of securities that are treated as separate "submarkets," but there can be logical dependencies among them.

Example 7. *Medal counts.* Consider a prediction market for the Olympics. Assume that Norway takes part in n Olympic events. In each, Norway can win a gold medal or not. Encode this outcome space as $\Omega = \{0, 1\}^n$. Define random variables $X_i(\boldsymbol{\omega}) = \omega_i$ equal to 1 iff Norway wins gold in the ith Olympic event. Also define a random variable $Y = \sum_{i=1}^n X_i$ representing the number of gold medals that Norway wins in total. We create $K = 2n + 1$ securities, corresponding to 0/1 indicators of the form $\mathbf{1}[X_i = 1]$ for $i \in [n]$ and $\mathbf{1}[Y = y]$ for $y \in \{0, \ldots, n\}$. That is, $\rho_i = X_i$ for $i \in [n]$ and $\rho_{n+1+y} = \mathbf{1}[Y = y]$ for $y \in \{0, \ldots, n\}$. A natural block structure in this market is $\mathcal{G} = \bigl\{\{1\}, \{2\}, \ldots, \{n\}, \{n + 1, \ldots, 2n + 1\}\bigr\}$ with submarkets corresponding to the X_i and Y.

Given the block structure $\mathcal{G}$, the construction of a linearly constrained market begins with bounded-loss and arbitrage-free convex cost functions $C_g : \mathbb{R}^g \to \mathbb{R}$ with conjugates R_g and divergences D_g for each $g \in \mathcal{G}$. These cost functions are assumed to be easy to compute and give rise to a "direct-sum" cost $C_\oplus(\boldsymbol{q}) = \sum_{g\in\mathcal{G}} C_g(\boldsymbol{q}_g)$ with the conjugate $R_\oplus(\boldsymbol{\mu}) = \sum_{g\in\mathcal{G}} R_g(\boldsymbol{\mu}_g)$ and divergence $D_\oplus(\boldsymbol{\mu}\|\boldsymbol{q}) = \sum_{g\in\mathcal{G}} D_g(\boldsymbol{\mu}_g\|\boldsymbol{q}_g)$.

Since $C_\oplus$ decomposes, it can be calculated quickly. However, the market maker $C_\oplus$ might allow arbitrage due to the lack of consistency among submarkets since arbitrage opportunities arise when prices fall outside $\mathcal{M}$ [1]. $\mathcal{M}$ is always polyhedral, so it can be described as $\mathcal{M} = \bigl\{\boldsymbol{\mu} \in \mathbb{R}^K : \mathbf{A}^\top\boldsymbol{\mu} \ge \boldsymbol{b}\bigr\}$ for some matrix $\mathbf{A} \in \mathbb{R}^{K\times M}$ and vector $\boldsymbol{b} \in \mathbb{R}^M$. Letting $\boldsymbol{a}_m$ denote the mth column of $\mathbf{A}$, arbitrage opportunities open up if the price of the bundle $\boldsymbol{a}_m$ falls below b_m. For any $\boldsymbol{\eta} \in \mathbb{R}^M_+$, the bundle $\mathbf{A}\boldsymbol{\eta}$ presents an arbitrage opportunity if priced below $\boldsymbol{b} \cdot \boldsymbol{\eta}$.

A *linearly constrained market maker* (LCMM) is described by the cost function $C(\boldsymbol{q}) = \inf_{\boldsymbol{\eta}\in\mathbb{R}^M_+}\bigl[C_\oplus(\boldsymbol{q}+\mathbf{A}\boldsymbol{\eta}) - \boldsymbol{b}\cdot\boldsymbol{\eta}\bigr]$. While the definition of C is slightly involved, the conjugate R has a natural meaning as a restriction of the direct-sum market to the price space $\mathcal{M}$, i.e., $R(\boldsymbol{\mu}) = R_{\oplus}(\boldsymbol{\mu}) + \mathbb{I}[\boldsymbol{\mu} \in \mathcal{M}]$. Furthermore, the infimum in the definition of C is always attained (see Appendix D.1). Fixing $\boldsymbol{q}$ and letting $\boldsymbol{\eta}^{\star}$ be a minimizer in the definition, we can think of the market maker as automatically charging traders for the bundle $\mathbf{A}\boldsymbol{\eta}^{\star}$, which would present an arbitrage opportunity, and returning to them the guaranteed payout $\boldsymbol{b} \cdot \boldsymbol{\eta}$. This benefits traders while maintaining the same worst-case loss guarantee for the market maker as $C_{\oplus}$ [13].

Example 8. *LCMM for medal counts.* Continuing the previous example, for submarkets X_i, we can define LMSR costs $C_i(q_i) = \ln(1 + \exp(q_i))$. For the submarket for Y, let $g = \{n+1, \dots, 2n+1\}$ and use the LMSR cost $C_g(\boldsymbol{q}_g) = \ln\bigl(\sum_{y=0}^{n} \exp(q_{n+1+y})\bigr)$. The submarkets for X_i and Y are linked. One example of a linear constraint is based on the linearity of expectations: for any distribution, we must have $\mathbb{E}[Y] = \sum_{i=1}^{n} \mathbb{E}[X_i]$. This places an equality constraint $\sum_{y=0}^{n} y \cdot \mu_{n+1+y} = \sum_{i=1}^{n} \mu_i$ on the vector $\boldsymbol{\mu}$, which can be expressed as two inequality constraints (see Dudík et al. [13, 14] for more on constraint generation).

4.2 DECREASING LIQUIDITY

We now study the gradual decrease scenario in which the utility for information in each submarket g decreases over time. In the Olympics example, the market maker may want to continuously decrease the rewards for information about a particular event as the event takes place.

We generalize the strategy from Example 6 to LCMMs and extend them to time-sensitive cost functions by introducing the "information-utility schedule" in the form of a differentiable non-increasing function $\beta_g : \mathbb{R} \to (0, 1]$ with $\beta_g(t^0) = 1$. The speed of decrease of β_g controls the speed of decrease of the utility for information in each submarket. (We make this statement more precise in Theorem 3.)

We first define a gradual decrease direct-sum cost function $\mathbf{C}_{\oplus}(\boldsymbol{q}; t) = \sum_{g \in \mathcal{G}} \beta_g(t) C_g\bigl(\boldsymbol{q}_g / \beta_g(t)\bigr)$ which is used to define a gradual decrease LCMM, and a matching `NewState` as follows:

$$\mathbf{C}(\boldsymbol{q}; t) = \inf_{\boldsymbol{\eta} \in \mathbb{R}^M_+} \bigl[\mathbf{C}_{\oplus}(\boldsymbol{q} + \mathbf{A}\boldsymbol{\eta}; t) - \boldsymbol{b} \cdot \boldsymbol{\eta}\bigr]$$

$$\texttt{NewState}(\boldsymbol{q};\, t, \tilde{t}) = \tilde{\boldsymbol{q}}$$

$$\text{such that } \tilde{\boldsymbol{q}}_g = \tfrac{\beta_g(\tilde{t})}{\beta_g(t)}(\boldsymbol{q}_g + \boldsymbol{\delta}^{\star}_g) - \boldsymbol{\delta}^{\star}_g$$

$$\text{where } \boldsymbol{\eta}^{\star} \text{ is a minimizer in } \mathbf{C}(\boldsymbol{q}; t) \text{ and } \boldsymbol{\delta}^{\star} = \mathbf{A}\boldsymbol{\eta}^{\star}\ .$$

When considering the state update from time t to time $\tilde{t}$, the ratio $\beta_g(\tilde{t})/\beta_g(t)$ has the role of the liquidity parameter α in Example 6. The motivation behind the definition of `NewState` is to guarantee that $\tilde{\boldsymbol{q}}_g + \boldsymbol{\delta}^{\star}_g = [\beta_g(\tilde{t})/\beta_g(t)](\boldsymbol{q}_g + \boldsymbol{\delta}^{\star}_g)$, which turns out to ensure that $\boldsymbol{\eta}^{\star}$ remains the minimizer and the prices are unchanged. The preservation of prices (PRICE) is achieved by a scaling similar to Example 6, albeit applied to the market state in the direct-sum market underlying the LCMM.

This intuition is formalized in the next theorem, which shows that the above construction preserves prices and decreases the utility for information, as captured by the mixed Bregman divergence, according to the schedules β_g. We use the notation $C^t(\boldsymbol{q}) := \mathbf{C}(\boldsymbol{q}; t)$ and write D^t_g for the divergence derived from $C^t_g(\boldsymbol{q}_g) := \beta_g(t) C_g(\boldsymbol{q}_g / \beta_g(t))$.

Theorem 3. *Let $\mathbf{C}$ be a gradual decrease LCMM, let $t, \tilde{t} \in \mathbb{R}$ and $\boldsymbol{s} \in \mathbb{R}^K$. The replacement of C^t by $\tilde{C} := C^{\tilde{t}}$ and $\boldsymbol{s}$ by $\tilde{\boldsymbol{s}} :=$ `NewState`$(\boldsymbol{s};\, t, \tilde{t})$ satisfies* PRICE. *Also,*

$$\tilde{D}(\boldsymbol{\mu} \| \tilde{\boldsymbol{s}}) = \sum_{g \in \mathcal{G}} \alpha_g D^t_g(\boldsymbol{\mu}_g \| \boldsymbol{s}_g + \boldsymbol{\delta}^{\star}_g) + (\mathbf{A}^{\top}\boldsymbol{\mu} - \boldsymbol{b}) \cdot \boldsymbol{\eta}^{\star} \quad (5)$$

for all $\boldsymbol{\mu} \in \mathcal{M}$, where $\boldsymbol{\eta}^{\star}$ and $\boldsymbol{\delta}^{\star}$ are defined by `NewState`$(\boldsymbol{s};\, t, \tilde{t})$*, and $\alpha_g = \beta_g(\tilde{t})/\beta_g(t) > 0$.*

The first term on the right-hand side of Eq. (5) is the sum of divergences in submarkets g, each weighted by a coefficient α_g which is equal to one at $\tilde{t} = t$ and weakly decreases as $\tilde{t}$ grows. The divergences are between $\boldsymbol{\mu}_g$ and the state resulting from the arbitrager action in the direct-sum market. The second term is non-negative, since $\boldsymbol{\mu} \in \mathcal{M}$, and represents expected arbitrager gains beyond the guaranteed profit from the arbitrage in the direct-sum market. The only terms that depend on time $\tilde{t}$ are the multipliers α_g. Since they are decreasing over time, we immediately obtain that the utility for information, `Util`$(\boldsymbol{\mu}; \tilde{\boldsymbol{s}}) = \tilde{D}(\boldsymbol{\mu} \| \tilde{\boldsymbol{s}})$, is also decreasing, with the contributions from individual submarkets decreasing according to their schedules β_g.

When only one of the schedules β_g is decreasing and the other schedules stay constant, we can show that the excess utility and conditional prices are preserved (conditioned on $\boldsymbol{\rho}_g$), and under certain conditions also DECUTIL holds.

For a submarket g, let $\mathcal{X}_g := \{\boldsymbol{\rho}_g(\omega) : \omega \in \Omega\}$ be the set of realizations of $\boldsymbol{\rho}_g$. Recall that $\mathcal{M}(\mathcal{E})$ is the convex hull of $\{\boldsymbol{\rho}(\omega)\}_{\omega \in \mathcal{E}}$. We show that DECUTIL holds if C_g is differentiable and the submarket g is "tight" as follows.

Definition 7. *We say that a submarket g is* tight *if for all $\boldsymbol{x} \in \mathcal{X}_g$ the set $\{\boldsymbol{\mu} \in \mathcal{M} : \boldsymbol{\mu}_g = \boldsymbol{x}\}$ coincides with $\mathcal{M}(\boldsymbol{\rho}_g = \boldsymbol{x})$, i.e., if all the beliefs $\boldsymbol{\mu}$ with $\boldsymbol{\mu}_g = \boldsymbol{x}$ can be realized by probability distributions over states ω with $\boldsymbol{\rho}_g(\omega) = \boldsymbol{x}$. (In general, the former is always a superset of the latter, hence the name "tight" when the equality holds.)*

While this condition is somewhat restrictive, it is easy to see that all submarkets with binary securities, i.e., with $\boldsymbol{\rho}_g(\omega) \in \{0, 1\}^g$, are tight (see Appendix D.4).

Theorem 4. *Assume the setup of Theorem 3. Let $g \in \mathcal{G}$ and assume that $\beta_g(\tilde{t}) < \beta_g(t)$ whereas $\beta_{g'}(\tilde{t}) = \beta_{g'}(t)$ for $g' \neq g$. Then the replacement of C^t by $\tilde{C}$ and $\boldsymbol{s}$ by $\tilde{\boldsymbol{s}}$ satisfies* CONDPRICE *and* EXUTIL *for the random variable $\boldsymbol{\rho}_g$. Furthermore, if C_g is differentiable and the submarket g is tight, we also obtain* DECUTIL.

References

[1] Jacob Abernethy, Yiling Chen, and Jennifer Wortman Vaughan. An optimization-based framework for automated market-making. In *Proceedings of the 12th ACM Conference on Electronic Commerce*, 2011.

[2] Jacob Abernethy, Yiling Chen, and Jennifer Wortman Vaughan. Efficient market making via convex optimization, and a connection to online learning. *ACM Transactions on Economics and Computation*, 1(2):12:1–12:39, 2013.

[3] Jacob Abernethy, Rafael Frongillo, Xiaolong Li, and Jennifer Wortman Vaughan. A general volume-parameterized market making framework. In *Proceedings of the 15th ACM Conference on Economics and Computation*, 2014.

[4] Joyce Berg, Robert Forsythe, Forrest Nelson, and Thomas Rietz. Results from a dozen years of election futures markets research. In Charles R. Plott and Vernon L. Smith, editors, *Handbook of Experimental Economics Results*, volume 1, pages 742–751. Elsevier, 2008.

[5] Aseem Brahma, Mithun Chakraborty, Sanmay Das, Allen Lavoie, and Malik Magdon-Ismail. A Bayesian market maker. In *Proceedings of the 13th ACM Conference on Electronic Commerce*, 2012.

[6] Philip Delves Broughton. Prediction markets: Value among the crowd. *Financial Times*, April 2013.

[7] Robert Charette. An internal futures market. *Information Management*, March 2007.

[8] Yiling Chen, Lance Fortnow, Evdokia Nikolova, and David M. Pennock. Betting on permutations. In *Proceedings of the 8th ACM Conference on Electronic Commerce*, 2007.

[9] Yiling Chen, Lance Fortnow, Nicolas Lambert, David M. Pennock, and Jennifer Wortman. Complexity of combinatorial market makers. In *Proceedings of the 9th ACM Conference on Electronic Commerce*, 2008.

[10] Yiling Chen, Sharad Goel, and David M. Pennock. Pricing combinatorial markets for tournaments. In *ACM Symposium on Theory of Computing*, 2008.

[11] Yiling Chen and David M. Pennock. A utility framework for bounded-loss market makers. In *Proceedings of the 23rd Conference on Uncertainty in Artificial Intelligence*, 2007.

[12] Bo Cowgill, Justin Wolfers, and Eric Zitzewitz. Using prediction markets to track information flows: Evidence from Google. Working paper, 2008.

[13] Miroslav Dudík, Sébastien Lahaie, and David M. Pennock. A tractable combinatorial market maker using constraint generation. In *Proceedings of the 13th ACM Conference on Electronic Commerce*, 2012.

[14] Miroslav Dudík, Sébastien Lahaie, David M. Pennock, and David Rothschild. A combinatorial prediction market for the U.S. elections. In *Proceedings of the 13th ACM Conference on Electronic Commerce*, 2013.

[15] Geoffrey J. Gordon. Regret bounds for prediction problems. In *Proceedings of the 12th Annual Conference on Computational Learning Theory*, pages 29–40, 1999.

[16] Peter D. Grünwald and A. Philip Dawid. Game theory, maximum entropy, minimum discrepancy and robust Bayesian decision theory. *The Annals of Statistics*, 32(4):1367–1433, 2004.

[17] Mingyu Guo and David M. Pennock. Combinatorial prediction markets for event hierarchies. In *International Conference on Autonomous Agents and Multiagent Systems*, 2009.

[18] Robin Hanson. Combinatorial information market design. *Information Systems Frontiers*, 5(1):105–119, 2003.

[19] Robin Hanson. Logarithmic market scoring rules for modular combinatorial information aggregation. *Journal of Prediction Markets*, 1(1):3–15, 2007.

[20] Jean-Baptiste Hiriart-Urruty and Claude Lemaréchal. *Convex Analysis and Minimization Algorithms*, volume 1. Springer, 1996.

[21] Xiaolong Li and Jennifer Wortman Vaughan. An axiomatic characterization of adaptive-liquidity market makers. In *Proceedings of the 14th ACM Conference on Electronic Commerce*, 2013.

[22] Abraham Othman and Tuomas Sandholm. Liquidity-sensitive automated market makers via homogeneous risk measures. In *Proceedings of the 7th Workshop on Internet and Network Economics*, 2011.

[23] Abraham Othman and Tuomas Sandholm. Profit-charging market makers with bounded loss, vanishing bid/ask spreads, and unlimited market depth. In *Proceedings of the 13th ACM Conference on Electronic Commerce*, 2012.

[24] Abraham Othman, Tuomas Sandholm, David M. Pennock, and Daniel M. Reeves. A practical liquidity-sensitive automated market maker. In *Proceedings of the 11th ACM Conference on Electronic Commerce*, 2010.

[25] David M. Pennock, Steve Lawrence, C. Lee Giles, and Finn A. Nielsen. The real power of artificial markets. *Science*, 291:987–988, 2002.

[26] David M. Pennock and Lirong Xia. Price updating in combinatorial prediction markets with Bayesian networks. In *Proceedings of the 27th Conference on Uncertainty in Artificial Intelligence*, 2011.

[27] R. Tyrrell Rockafellar. *Convex Analysis*. Princeton University Press, 1970.

[28] Martin Spann and Bernd Skiera. Internet-based virtual stock markets for business forecasting. *Management Science*, 49 (10):1310–1326, 2003.

[29] Lyle Ungar, Barb Mellors, Ville Satopää, Jon Baron, Phil Tetlock, Jaime Ramos, and Sam Swift. The good judgment project: A large scale test of different methods of combining expert predictions. AAAI Technical Report FS-12-06, 2012.

[30] Jean-Baptiste Hiriart-Urruty and Claude Lemaréchal. *Fundamentals of Convex Analysis*. Springer, 2001.

A CONVEX ANALYSIS

Here we briefly review concepts and results from convex analysis which we use throughout the paper.

Convex sets, polytopes, relative interior. Let $S \subseteq \mathbb{R}^n$. We say that S is *convex* if it contains all line segments with endpoints in S. The *convex hull* of S, denoted $\operatorname{conv} S$, is the smallest convex set containing S. It can be characterized as the set containing all "convex combinations" of points in S [27, Theorem 2.3], where a *convex combination* of points $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k$ is a point $\boldsymbol{u} = \sum_{i=1}^{k} \lambda_i \boldsymbol{u}_i$ for any $\lambda_i \geq 0$ with $\sum_{i=1}^{k} \lambda_i = 1$.

A set which is a convex hull of a finite set of points is called a *polytope*. We say that S is *polyhedral* if it is an intersection of a finite set of half-spaces, i.e., if $S = \{\boldsymbol{u} \in \mathbb{R}^n : \mathbf{A}\boldsymbol{u} \geq \boldsymbol{b}\}$ for some matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ and vector $\boldsymbol{b} \in \mathbb{R}^m$. All polytopes are polyhedral [27, Theorem 19.1].

An *affine hull* of S is the smallest affine space containing S. The topological interior of S relative to its affine hull is called the *relative interior* of S and denoted $\operatorname{relint} S$. To give a common example, if S is a simplex in n dimensions, i.e., $S = \{\boldsymbol{u} \in \mathbb{R}^n : u_i \geq 0, \sum_{i=1}^{n} u_i = 1\}$, then the interior of S is empty, but $\operatorname{relint} S = \{\boldsymbol{u} \in \mathbb{R}^n : u_i > 0, \sum_{i=1}^{n} u_i = 1\}$.

Function properties, epigraph, closure, roof. Consider a function $f : \mathbb{R}^n \to (-\infty, \infty]$. Its *domain*, denoted $\operatorname{dom} f$, is the set of points $\boldsymbol{u}$ such that $f(\boldsymbol{u})$ is finite. The function f is called *proper* if its domain is non-empty. The *epigraph* of f, denoted $\operatorname{epi} f$, is the set of points on and above the graph of f, i.e.,

$$\operatorname{epi} f := \{(\boldsymbol{u}, t) \in \mathbb{R}^n \times \mathbb{R} : \ t \geq f(\boldsymbol{u})\} \ .$$

The function f is called *closed* if its epigraph is a closed set. This is equivalent to f being lower semi-continuous [27, Theorem 7.1]. The function f is called *convex* if its epigraph is a convex set, or equivalently, if for all $\boldsymbol{u}, \boldsymbol{u}' \in \operatorname{dom} f$, for all $\lambda \in (0, 1)$,

$$f(\lambda \boldsymbol{u} + (1 - \lambda)\boldsymbol{u}') \leq \lambda f(\boldsymbol{u}) + (1 - \lambda) f(\boldsymbol{u}') \ .$$

The function f is *strictly convex* if the inequality above is strict whenever $\boldsymbol{u} \neq \boldsymbol{u}'$. Closed convex functions are not only lower semi-continuous, but actually continuous relative to any polyhedral subset of their domain (see Theorems 10.2 and 20.5 of Rockafellar [27]).

Proposition 3. *Let $f : \mathbb{R}^n \to (-\infty, \infty]$ be a closed convex function and S any polyhedral subset of $\operatorname{dom} f$. Then S is continuous relative to S.*

Any convex function finite on all of $\mathbb{R}^n$ is necessarily continuous [27, Corollary 10.1.1] and therefore closed. The *closure* of f, denoted $\operatorname{cl} f$, is the unique function whose epigraph is the topological closure of $\operatorname{epi} f$. As defined in Sec. 3, the *convex roof* of f, denoted $(\operatorname{conv} f)$, is the unique function whose epigraph is the convex hull of $\operatorname{epi} f$.

Subdifferential, conjugacy, duality. Consider a convex function $f : \mathbb{R}^n \to (-\infty, \infty]$. A *subgradient* of f at a point $\boldsymbol{u} \in \operatorname{dom} f$ is a vector $\boldsymbol{v} \in \mathbb{R}^n$ such that

$$f(\boldsymbol{u}') \geq f(\boldsymbol{u}) + \boldsymbol{v} \cdot (\boldsymbol{u}' - \boldsymbol{u})$$

for all $\boldsymbol{u}'$. The set of all subgradients of f at $\boldsymbol{u}$ is called the *subdifferential* of f at $\boldsymbol{u}$ and denoted $\partial f(\boldsymbol{u})$. If f is differentiable at $\boldsymbol{u}$, then $\partial f(\boldsymbol{u})$ is the singleton equal to the gradient of f at $\boldsymbol{u}$.

Let $f : \mathbb{R}^n \to (-\infty, \infty]$ be any proper function. The *(convex) conjugate* of f is the function $f^* : \mathbb{R}^n \to (-\infty, \infty]$ defined by

$$f^*(\boldsymbol{v}) := \sup_{\boldsymbol{u} \in \mathbb{R}^n} [\boldsymbol{v} \cdot \boldsymbol{u} - f(\boldsymbol{u})] \ . \tag{6}$$

The function f^* is always closed and convex (because its epigraph is an intersection of half-spaces). We write $f^{**} = (f^*)^*$ to denote the *biconjugate* of f. The biconjugate is a closure of the convex roof of f [20, Theorem E.1.3.5].

Proposition 4. *Let $f : \mathbb{R}^n \to (-\infty, \infty]$ be a proper convex function or a proper function bounded below by an affine function. Then $f^{**} = \operatorname{cl}(\operatorname{conv} f)$. Hence, if f is a closed proper convex function, $f^{**} = f$.*

The definition of the conjugate implies that

$$f^*(\boldsymbol{v}) \geq \boldsymbol{v} \cdot \boldsymbol{u} - f(\boldsymbol{u}) \tag{7}$$

for all $\boldsymbol{u}$ and $\boldsymbol{v}$, with the equality if and only if $\boldsymbol{u}$ is the maximizer on the right-hand side of Eq. (6). If f is convex, this can only happen if $\boldsymbol{v} \in \partial f(\boldsymbol{u})$. Similar reasoning can be applied to f^{**}, yielding the following proposition (based on Theorem 23.5 of Rockafellar [27]). Instead of f and f^*, we use the notation C and R to reflect the intended use in the body of the paper. The gap between the left-hand side and the right-hand side of Eq. (7) is referred to as the *mixed Bregman divergence*.

Proposition 5. *Let C be a closed proper convex function, R its conjugate, and D the associated mixed Bregman divergence $D(\boldsymbol{\mu} \| \boldsymbol{q}) := R(\boldsymbol{\mu}) + C(\boldsymbol{q}) - \boldsymbol{\mu} \cdot \boldsymbol{q}$. Then $D(\boldsymbol{\mu} \| \boldsymbol{q}) \geq 0$ for all $\boldsymbol{\mu}$, $\boldsymbol{q}$ and the following statements are equivalent:*

- $D(\boldsymbol{\mu} \| \boldsymbol{q}) = 0$
- $\boldsymbol{q} \in \partial R(\boldsymbol{\mu})$
- $\boldsymbol{\mu} \in \partial C(\boldsymbol{q})$

A function is called *polyhedral* if its epigraph is polyhedral. The following theorem relates a convex minimization problem with a concave maximization problem via convex conjugates. It is a version of *Fenchel's duality* and a subcase of Corollary 31.2.1 of Rockafellar [27].

Theorem 5 (Fenchel's duality). *Let $f : \mathbb{R}^K \to (-\infty, \infty]$ and $g : \mathbb{R}^M \to (-\infty, \infty]$ be closed convex functions and $\mathbf{A} \in \mathbb{R}^{K\times M}$. Further assume that g is polyhedral and there exists $\boldsymbol{\mu} \in \operatorname{relint}(\operatorname{dom} f^*)$ such that $\mathbf{A}^\top\boldsymbol{\mu} \in \operatorname{dom} g^*$. Then*

$$\inf_{\boldsymbol{\eta}\in\mathbb{R}^M} [f(\mathbf{A}\boldsymbol{\eta}) + g(\boldsymbol{\eta})] = \sup_{\boldsymbol{\mu}\in\mathbb{R}^K} \left[-f^*(\boldsymbol{\mu}) - g^*(-\mathbf{A}^\top\boldsymbol{\mu})\right]$$

and the infimum is attained.

B PROOFS FROM SECTION 2.5

B.1 PROOF OF THEOREM 1

First we prove Eq. (1) using the definition of $\texttt{Util}(\boldsymbol{\mu};\boldsymbol{q})$ and the conjugacy of R and C:

$$\begin{aligned}\texttt{Util}(\boldsymbol{\mu};\boldsymbol{q}) &= \sup\nolimits_{\boldsymbol{r}\in\mathbb{R}^K}\left[\boldsymbol{\mu}\cdot\boldsymbol{r} - C(\boldsymbol{q}+\boldsymbol{r}) + C(\boldsymbol{q})\right]\\ &= \sup\nolimits_{\boldsymbol{r}\in\mathbb{R}^K}\left[\boldsymbol{\mu}\cdot(\boldsymbol{q}+\boldsymbol{r}) - C(\boldsymbol{q}+\boldsymbol{r}) - \boldsymbol{\mu}\cdot\boldsymbol{q} + C(\boldsymbol{q})\right]\\ &= R(\boldsymbol{\mu}) - \boldsymbol{\mu}\cdot\boldsymbol{q} + C(\boldsymbol{q}) = D(\boldsymbol{\mu}\|\boldsymbol{q}) \ .\end{aligned}$$

Next, we prove Eq. (2):

$$\begin{aligned}\texttt{Util}(\mathcal{E};\boldsymbol{q}) &= \sup_{\boldsymbol{r}\in\mathbb{R}^K}\min_{\omega\in\mathcal{E}}\Big[\boldsymbol{r}\cdot\boldsymbol{\rho}(\omega) + C(\boldsymbol{q}) - C(\boldsymbol{q}+\boldsymbol{r})\Big]\\ &= \sup_{\boldsymbol{q}'\in\mathbb{R}^K}\min_{\boldsymbol{\mu}'\in\mathcal{M}(\mathcal{E})}\Big[(\boldsymbol{q}'-\boldsymbol{q})\cdot\boldsymbol{\mu}' + C(\boldsymbol{q}) - C(\boldsymbol{q}')\Big] \qquad (8)\\ &= \min_{\boldsymbol{\mu}'\in\mathcal{M}(\mathcal{E})}\sup_{\boldsymbol{q}'\in\mathbb{R}^K}\Big[(\boldsymbol{q}'-\boldsymbol{q})\cdot\boldsymbol{\mu}' + C(\boldsymbol{q}) - C(\boldsymbol{q}')\Big] \qquad (9)\\ &= \min_{\boldsymbol{\mu}'\in\mathcal{M}(\mathcal{E})}\Big[R(\boldsymbol{\mu}') - \boldsymbol{q}\cdot\boldsymbol{\mu}' + C(\boldsymbol{q})\Big] \qquad (10)\\ &= \min_{\boldsymbol{\mu}'\in\mathcal{M}(\mathcal{E})} D(\boldsymbol{\mu}'\|\boldsymbol{q}) \ ,\end{aligned}$$

where the equalities are justified as follows. Eq. (8) follows by relaxing, without loss of generality, the optimization of a linear function over $\{\boldsymbol{\rho}(\omega)\}_{\omega\in\mathcal{E}}$ to the optimization over the convex hull, and substituting $\boldsymbol{q}' = \boldsymbol{q}+\boldsymbol{r}$. Eq. (9) follows from Sion's minimax theorem, and finally Eq. (10) follows from the definition of convex conjugacy.

The final statement to prove, Eq. (3), follows immediately from the definition of $\boldsymbol{p}(\mathcal{E};\boldsymbol{q})$ and Eqs. (1) and (2).

B.2 PROOF OF PROPOSITION 1

It suffices to show that the statement of the proposition holds for a specific $x \in \mathcal{X}$ with ExUtil and CondPrice also restricted to a specific x. The proposition will then follow by universal quantification across all $x \in \mathcal{X}$. Thus in the remainder we consider a specific $x \in \mathcal{X}$.

Let $\hat{\boldsymbol{\mu}}^x \in \boldsymbol{p}(X = x;\boldsymbol{s})$ and $\tilde{\boldsymbol{\mu}}^x \in \tilde{\boldsymbol{p}}(X = x;\tilde{\boldsymbol{s}})$. The definition of the excess utility for a belief and Theorem 1 then imply that ExUtil (restricted to x) is satisfied if and only if for all $\boldsymbol{\mu} \in \mathcal{M}^x$

$$D(\boldsymbol{\mu}\|\boldsymbol{s}) - D(\hat{\boldsymbol{\mu}}^x\|\boldsymbol{s}) = \tilde{D}(\boldsymbol{\mu}\|\tilde{\boldsymbol{s}}) - \tilde{D}(\tilde{\boldsymbol{\mu}}^x\|\tilde{\boldsymbol{s}}) \ . \qquad (11)$$

First assume that ExUtil holds and therefore Eq. (11) holds for x. Then the desired condition follows by setting $c^x = D(\hat{\boldsymbol{\mu}}^x\|\boldsymbol{s}) - \tilde{D}(\tilde{\boldsymbol{\mu}}^x\|\tilde{\boldsymbol{s}})$.

Conversely, assume that $D(\boldsymbol{\mu}\|\boldsymbol{s}) - \tilde{D}(\boldsymbol{\mu}\|\tilde{\boldsymbol{s}}) = c^x$ holds for all $\boldsymbol{\mu} \in \mathcal{M}^x$. Since $D(\boldsymbol{\mu}\|\boldsymbol{s}) = \tilde{D}(\boldsymbol{\mu}\|\tilde{\boldsymbol{s}}) + c^x$, we obtain

$$\begin{aligned}\boldsymbol{p}(X = x;\boldsymbol{s}) &= \operatorname*{argmin}_{\boldsymbol{\mu}\in\mathcal{M}^x} D(\boldsymbol{\mu}\|\boldsymbol{s})\\ &= \operatorname*{argmin}_{\boldsymbol{\mu}\in\mathcal{M}^x} \tilde{D}(\boldsymbol{\mu}\|\tilde{\boldsymbol{s}}) = \tilde{\boldsymbol{p}}(X = x;\tilde{\boldsymbol{s}}) \ ,\end{aligned}$$

i.e., CondPrice (restricted to x) holds. This and the argument above show that ExUtil implies CondPrice.

To finish the proof we have to show that the assumption that $D(\boldsymbol{\mu}\|\boldsymbol{s}) - \tilde{D}(\boldsymbol{\mu}\|\tilde{\boldsymbol{s}}) = c^x$ also implies ExUtil. Let $\hat{\boldsymbol{\mu}}^x \in \boldsymbol{p}(X = x;\boldsymbol{s}) = \tilde{\boldsymbol{p}}(X = x;\tilde{\boldsymbol{s}})$. Then we have

$$D(\boldsymbol{\mu}\|\boldsymbol{s}) - \tilde{D}(\boldsymbol{\mu}\|\tilde{\boldsymbol{s}}) = c^x = D(\hat{\boldsymbol{\mu}}^x\|\boldsymbol{s}) - \tilde{D}(\hat{\boldsymbol{\mu}}^x\|\tilde{\boldsymbol{s}})$$

and rearranging yields Eq. (11), with $\hat{\boldsymbol{\mu}}^x$ substituted for $\tilde{\boldsymbol{\mu}}^x$. However, $\hat{\boldsymbol{\mu}}^x$ is a valid choice of $\tilde{\boldsymbol{\mu}}^x$ since $\tilde{\boldsymbol{\mu}}^x$ was chosen arbitrarily from $\tilde{\boldsymbol{p}}(X = x;\tilde{\boldsymbol{s}}) = \boldsymbol{p}(X = x;\boldsymbol{s})$, so Eq. (11) and therefore ExUtil hold.

C PROOFS FROM SECTION 3

C.1 PROOF OF LEMMA 1

Theorem 1 shows that all the desiderata, except for Price, are derived from properties of $\tilde{D}(\boldsymbol{\mu}\|\tilde{\boldsymbol{s}})$ as a function of $\boldsymbol{\mu}$. To see that Price can also be derived this way, note that by Prop. 5 we have $\tilde{\boldsymbol{p}}(\tilde{\boldsymbol{s}}) = \partial\tilde{C}(\tilde{\boldsymbol{s}}) = \{\boldsymbol{\mu} :\ \tilde{D}(\boldsymbol{\mu}\|\tilde{\boldsymbol{s}}) = 0\}$. Thus, it suffices to analyze $\tilde{D}$. With $\tilde{C}'$ and $\tilde{\boldsymbol{s}}'$ as in the lemma, we have $\tilde{R}'(\boldsymbol{\mu}) = \tilde{R}(\boldsymbol{\mu}) - (\tilde{\boldsymbol{s}} - \boldsymbol{s})\cdot\boldsymbol{\mu}$ and

$$\tilde{D}'(\boldsymbol{\mu}\|\boldsymbol{s}) = \tilde{R}(\boldsymbol{\mu}) - (\tilde{\boldsymbol{s}}-\boldsymbol{s})\cdot\boldsymbol{\mu} + \tilde{C}(\tilde{\boldsymbol{s}}) - \boldsymbol{\mu}\cdot\boldsymbol{s} = \tilde{D}(\boldsymbol{\mu}\|\tilde{\boldsymbol{s}}).$$

Hence the lemma holds.

C.2 PROOF OF LEMMA 2

Prop. 1 shows that ExUtil and CondPrice are together satisfied if and only if there exist constants c^x such that for all $x \in \mathcal{X}$ and $\boldsymbol{\mu} \in \mathcal{M}^x$,

$$c^x = D(\boldsymbol{\mu}\|\boldsymbol{s}) - \tilde{D}(\boldsymbol{\mu}\|\boldsymbol{s}) = C(\boldsymbol{s}) + R(\boldsymbol{\mu}) - \tilde{C}(\boldsymbol{s}) - \tilde{R}(\boldsymbol{\mu}) \ .$$

If this statement holds, then for any x, setting $b^x = c^x - C(\boldsymbol{s}) + \tilde{C}(\boldsymbol{s})$ gives us $\tilde{R}(\boldsymbol{\mu}) = R(\boldsymbol{\mu}) - b^x$ for all $\boldsymbol{\mu} \in \mathcal{M}^x$. Conversely, if $\tilde{R}(\boldsymbol{\mu}) = R(\boldsymbol{\mu}) - b^x$ for all $\boldsymbol{\mu} \in \mathcal{M}^x$, setting $c^x = b^x + C(\boldsymbol{s}) - \tilde{C}(\boldsymbol{s})$ gives the equation above.

C.3 PROOF OF LEMMA 3

From the definition of convex roof, we have

$$\tilde{R}'(\boldsymbol{\mu}) = \sup\left\{g(\boldsymbol{\mu}) :\ g \in \mathcal{G},\ g \le R^{\boldsymbol{b}}\right\} \ . \qquad (12)$$

Since we have $\tilde{R} \leq R^{\boldsymbol{b}}$, the function $\tilde{R}$ is a valid choice for g in Eq. (12). This gives us $\tilde{R} \leq \tilde{R}' \leq R^{\boldsymbol{b}}$ and thus $\tilde{R}'$ must be consistent with $R^{\boldsymbol{b}}$, proving the first part.

To prove the second part, we show a stronger statement:

$$\tilde{D}'(\boldsymbol{\mu}\|\boldsymbol{q}) \leq \tilde{D}(\boldsymbol{\mu}\|\boldsymbol{q}) \text{ for all } \boldsymbol{\mu} \in \mathcal{M}^\star, \boldsymbol{q} \in \mathbb{R}^K, \quad (13)$$

where $\tilde{D}'$ is the mixed Bregman divergence with respect to $\tilde{R}'$. The second part follows from Eq. (13) by setting $\boldsymbol{q} = \boldsymbol{s}$ and $\boldsymbol{\mu} = \hat{\boldsymbol{\mu}}^x$ (for ZEROUTIL), or choosing arbitrary $\boldsymbol{\mu} \in \mathcal{M}^\star$ (for DECUTIL). It remains to prove Eq. (13).

Since $\tilde{R}' \geq \tilde{R}$, we have for their conjugates $\tilde{C}' \leq \tilde{C}$. Also, for any $\boldsymbol{\mu} \in \mathcal{M}^\star$ we have $\tilde{R}'(\boldsymbol{\mu}) = R^{\boldsymbol{b}}(\boldsymbol{\mu}) = \tilde{R}(\boldsymbol{\mu})$, and thus

$$\begin{aligned}\tilde{D}'(\boldsymbol{\mu}\|\boldsymbol{q}) &= \tilde{R}'(\boldsymbol{\mu}) + \tilde{C}'(\boldsymbol{q}) - \boldsymbol{q}\cdot\boldsymbol{\mu} \\ &\leq \tilde{R}(\boldsymbol{\mu}) + \tilde{C}(\boldsymbol{q}) - \boldsymbol{q}\cdot\boldsymbol{\mu} = \tilde{D}(\boldsymbol{\mu}\|\boldsymbol{q}) \ . \qquad \square\end{aligned}$$

C.4 PROOF OF THEOREM 2

First, assume that CONDPRICE, EXUTIL, and ZEROUTIL are simultaneously satisfiable using Protocol 1. By Lemma 1, this implies they are satisfiable with the identity function for `NewState` and some function `NewCost`.

By Lemmas 2 and 3, it must be the case that for any state $\boldsymbol{s}$, there exists some $\boldsymbol{b} \in \mathbb{R}^{\mathcal{X}}$, such that the conditions would remain satisfied if `NewCost`$(\boldsymbol{s})$ instead output the conjugate $\tilde{C}$ of $\tilde{R} := (\operatorname{conv} R^{\boldsymbol{b}})$. It remains to show that the three conditions would remain satisfied if `NewCost`$(\boldsymbol{s})$ output the conjugate of $(\operatorname{conv} R^{\hat{\boldsymbol{b}}})$.

For all $x \in \mathcal{X}$, we can simplify $\tilde{\texttt{Util}}(X = x; \boldsymbol{s})$ using Eq. (2) and Eq. (3) as follows:

$$\begin{aligned}\tilde{\texttt{Util}}(X = x; \boldsymbol{s}) &= \tilde{D}(\hat{\boldsymbol{\mu}}^x\|\boldsymbol{s}) \\ &= \tilde{C}(\boldsymbol{s}) + \tilde{R}(\hat{\boldsymbol{\mu}}^x) - \boldsymbol{s}\cdot\hat{\boldsymbol{\mu}}^x \\ &= \tilde{C}(\boldsymbol{s}) + R(\hat{\boldsymbol{\mu}}^x) - b^x - \boldsymbol{s}\cdot\hat{\boldsymbol{\mu}}^x \quad (14) \\ &= \tilde{C}(\boldsymbol{s}) - C(\boldsymbol{s}) + \hat{b}^x - b^x \quad (15)\end{aligned}$$

for some $\hat{\boldsymbol{\mu}}^x \in \boldsymbol{p}(\Omega^x; \boldsymbol{s})$. Here Eq. (14) follows by consistency of $\tilde{R}$ with $R^{\boldsymbol{b}}$ and Eq. (15) follows by the definition of $\hat{b}^x$ in Eq. (4). Since ZEROUTIL is satisfied, $\tilde{\texttt{Util}}(X = x; \boldsymbol{s}) = 0$ for all $x \in \mathcal{X}$, so $\tilde{D}(\hat{\boldsymbol{\mu}}^x\|\boldsymbol{s}) = 0 = \tilde{D}(\hat{\boldsymbol{\mu}}^{x'}\|\boldsymbol{s})$ for all $x, x' \in \mathcal{X}$. Eq. (15) then yields

$$\tilde{C}(\boldsymbol{s}) - C(\boldsymbol{s}) + \hat{b}^x - b^x = \tilde{C}(\boldsymbol{s}) - C(\boldsymbol{s}) + \hat{b}^{x'} - b^{x'} \ .$$

Canceling the constant terms $\tilde{C}(\boldsymbol{s})$ and $C(\boldsymbol{s})$, we obtain that $\boldsymbol{b} = \hat{\boldsymbol{b}} + c\mathbf{1}$ for some $c \in \mathbb{R}$. Since $R^{\boldsymbol{b}} = R^{\hat{\boldsymbol{b}}} - c$, and $\tilde{R} = (\operatorname{conv} R^{\boldsymbol{b}})$ is consistent with $R^{\boldsymbol{b}}$, we conclude that $(\operatorname{conv} R^{\hat{\boldsymbol{b}}}) = (\operatorname{conv} R^{\boldsymbol{b}}) - c$ is consistent with $R^{\hat{\boldsymbol{b}}}$, and therefore by Lemma 2, EXUTIL and CONDPRICE remain satisfied switching to $(\operatorname{conv} R^{\hat{\boldsymbol{b}}})$. Additionally, since $(\operatorname{conv} R^{\boldsymbol{b}})$ and $(\operatorname{conv} R^{\hat{\boldsymbol{b}}})$ differ only by a vertical shift, the divergences associated with both are identical, and ZEROUTIL is also satisfied.

For the converse, assume $\tilde{R} := (\operatorname{conv} R^{\hat{\boldsymbol{b}}})$ is consistent with $R^{\hat{\boldsymbol{b}}}$. By Lemma 2, EXUTIL and CONDPRICE are satisfied, and it remains only to show $\tilde{\texttt{Util}}(X = x; \boldsymbol{s}) = 0$. This follows from Eq. (15), since now $\boldsymbol{b} = \hat{\boldsymbol{b}}$, and by Prop. 2, $\tilde{C}(\boldsymbol{s}) = C(\boldsymbol{s})$. (Note that Prop. 2 is stated after Theorem 2 in the main text, but its proof, given in the next section, does not rely on Theorem 2.)

C.5 PROOF OF PROPOSITION 2

We first show that $(\operatorname{conv} R^{\hat{\boldsymbol{b}}})$ is closed. Since R is the conjugate of C, it must be closed (see Appendix A). The domain of R is $\mathcal{M}$ which is polyhedral, and therefore R is in fact continuous on $\mathcal{M}$ (by Prop. 3). Since $\mathcal{M}$ is compact, R attains a maximum on $\mathcal{M}$, and in particular is bounded above on $\mathcal{M}$. Thus, also $R^{\hat{\boldsymbol{b}}}$ is bounded above on $\mathcal{M}^\star$. Let $u \in \mathbb{R}$ be the corresponding upper bound, i.e., $R^{\hat{\boldsymbol{b}}}(\boldsymbol{\mu}) \leq u$ for all $\boldsymbol{\mu} \in \mathcal{M}^\star$. We may write $\operatorname{epi} \tilde{R} = \operatorname{conv}(\operatorname{epi} R^{\hat{\boldsymbol{b}}})$ by definition of the roof construction. Now we can chop off $\operatorname{epi} R^{\hat{\boldsymbol{b}}}$ at u and consider the remainder:

$$\begin{aligned}S &= \{(\boldsymbol{\mu}, t) : \boldsymbol{\mu} \in \mathcal{M}^\star,\ R^{\hat{\boldsymbol{b}}}(\boldsymbol{\mu}) \leq t \leq u\} \\ &= \bigcup_{x\in\mathcal{X}} \{(\boldsymbol{\mu}, t) : \boldsymbol{\mu} \in \mathcal{M}^x,\ R^{\hat{\boldsymbol{b}}}(\boldsymbol{x}) \leq t \leq u\} \ . \quad (16)\end{aligned}$$

The set S is compact, because it is a finite union of compact sets in Eq. (16). Each individual term in Eq. (16) is indeed compact, because it is bounded (above by u and below by the boundedness of R on $\mathcal{M}^x$) and closed (by closedness of R and closedness of $\mathcal{M}^x$). Since S is compact, $\operatorname{conv} S$ is closed. Therefore,

$$\operatorname{epi} \tilde{R} = \operatorname{conv}(\operatorname{epi} R^{\hat{\boldsymbol{b}}}) = (\operatorname{conv} S) \cup \big(\mathcal{M} \times [u, \infty)\big)$$

is also a closed set, and thus $\tilde{R}$ is a closed convex function.

Recall from standard convex analysis (see Appendix A) that for any function f, we have $f^{**} = \operatorname{cl}(\operatorname{conv} f)$; the biconjugate of f is the closed convex roof of f. As we have shown, $(\operatorname{conv} R^{\hat{\boldsymbol{b}}}) = \operatorname{cl}(\operatorname{conv} R^{\hat{\boldsymbol{b}}}) = (R^{\hat{\boldsymbol{b}}})^{**}$, so $\tilde{R} = (R^{\hat{\boldsymbol{b}}})^{**}$, and in particular, $\tilde{C} = \tilde{R}^* = (R^{\hat{\boldsymbol{b}}})^{***} = (R^{\hat{\boldsymbol{b}}})^*$. Now, calculate

$$\begin{aligned}\tilde{C}(\boldsymbol{q}) &= \sup_{\boldsymbol{\mu}\in\mathcal{M}^\star} \Big[\boldsymbol{q}\cdot\boldsymbol{\mu} - R^{\hat{\boldsymbol{b}}}(\boldsymbol{\mu})\Big] \\ &= \max_{x\in\mathcal{X}} \sup_{\boldsymbol{\mu}\in\mathcal{M}^x} \Big[\boldsymbol{q}\cdot\boldsymbol{\mu} - R(\boldsymbol{\mu}) + \hat{b}^x\Big] \\ &= \max_{x\in\mathcal{X}} \Big[\hat{b}^x + \sup_{\boldsymbol{\mu}\in\mathcal{M}^x} \big[\boldsymbol{q}\cdot\boldsymbol{\mu} - R(\boldsymbol{\mu})\big]\Big] \\ &= \max_{x\in\mathcal{X}} \Big[\hat{b}^x + C^x(\boldsymbol{q})\Big] \ .\end{aligned}$$

Finally, observe that by definition of $\hat{b}^x$ and Theorem 1,

$$\hat{b}^x = C(\boldsymbol{s}) - \sup_{\boldsymbol{\mu}\in\mathcal{M}^x} \big[\boldsymbol{\mu}\cdot\boldsymbol{s} - R(\boldsymbol{\mu})\big] = C(\boldsymbol{s}) - C^x(\boldsymbol{s}) \ .$$

D PROOFS FROM SECTION 4

D.1 PROPERTIES OF LCMMS

The following properties of LCMM are used in the sequel.

Theorem 6. *Let C be a linearly constrained market maker with*

$$C(\boldsymbol{q}) = \inf_{\boldsymbol{\eta}\in\mathbb{R}^M_+}\left[C_\oplus(\boldsymbol{q}+\mathbf{A}\boldsymbol{\eta}) - \boldsymbol{b}\cdot\boldsymbol{\eta}\right] \ . \tag{17}$$

It has the following properties:

(a) *The conjugate R is a restriction of $R_\oplus$ to $\mathcal{M}$:*

$$R(\boldsymbol{\mu}) = R_\oplus(\boldsymbol{\mu}) + \mathbb{I}\left[\boldsymbol{\mu}\in\mathcal{M}\right] \ .$$

(b) *For every $\boldsymbol{q}$, there exists a minimizer $\boldsymbol{\eta}^\star$ of Eq. (17).*

(c) *Let $\boldsymbol{\eta}^\star$ be a minimizer of Eq. (17) for a specific $\boldsymbol{q}$ and let $\boldsymbol{\delta}^\star = \mathbf{A}\boldsymbol{\eta}^\star$. The Bregman divergence from $\boldsymbol{q}$ is then*

$$D(\boldsymbol{\mu}\|\boldsymbol{q}) = D_\oplus(\boldsymbol{\mu}\|\boldsymbol{q}+\boldsymbol{\delta}^\star) + (\mathbf{A}^\top\boldsymbol{\mu}-\boldsymbol{b})\cdot\boldsymbol{\eta}^\star + \mathbb{I}[\boldsymbol{\mu}\in\mathcal{M}] \, .$$

(d) *Let $\boldsymbol{\eta} \geq \mathbf{0}$ and $\boldsymbol{\delta} = \mathbf{A}\boldsymbol{\eta}$. Then $\boldsymbol{\eta}$ is a minimizer of Eq. (17) for a specific $\boldsymbol{q}$ if and only if there exists some $\boldsymbol{\mu}\in\mathcal{M}$ such that*

$$D_\oplus(\boldsymbol{\mu}\|\boldsymbol{q}+\boldsymbol{\delta}) + (\mathbf{A}^\top\boldsymbol{\mu}-\boldsymbol{b})\cdot\boldsymbol{\eta} = 0 \ .$$

Part (a) shows that while the definition of C in Eq. (17) is slightly involved, the conjugate R has a natural meaning as a restriction of the direct-sum market to the price space $\mathcal{M}$. Part (b) shows that we can take the minimum rather than the infimum in the definition of C, i.e., there is an optimal arbitrage bundle. Part (c) decomposes the Bregman divergence (and thus utility for information) into three terms. The last term forces $\boldsymbol{\mu}\in\mathcal{M}$. The first term is the (direct-sum) divergence between $\boldsymbol{\mu}$ and the state resulting from the arbitrager action in the direct-sum market. The second term is non-negative for $\boldsymbol{\mu}\in\mathcal{M}$, and represents expected arbitrager gains beyond the guaranteed profit from the arbitrage. Part (d) spells out first-order optimality conditions for an optimal arbitrage bundle $\boldsymbol{\eta}$.

Proof. We prove the theorem in parts.

Parts (a) and (b) We use a version of Fenchel's duality from Theorem 5. Specifically, consider a fixed $\boldsymbol{q}\in\mathbb{R}^K$ and let f and g be defined by

$$f(\boldsymbol{u}) = C_\oplus(\boldsymbol{q}+\boldsymbol{u}) \ , \quad g(\boldsymbol{\eta}) = \mathbb{I}[\boldsymbol{\eta}\geq\mathbf{0}] - \boldsymbol{b}\cdot\boldsymbol{\eta}$$

and hence their conjugates are

$$f^*(\boldsymbol{\mu}) = R_\oplus(\boldsymbol{\mu}) - \boldsymbol{q}\cdot\boldsymbol{\mu} \ , \quad g^*(\boldsymbol{v}) = \mathbb{I}[\boldsymbol{v}+\boldsymbol{b}\leq\mathbf{0}] \ .$$

Assuming that the conditions of Theorem 5 are satisfied for f and g, and plugging in the above definitions, we obtain

$$\begin{aligned}C(\boldsymbol{q}) &= \inf_{\boldsymbol{\eta}\in\mathbb{R}^M}\Big[C_\oplus(\boldsymbol{q}+\mathbf{A}\boldsymbol{\eta}) - \boldsymbol{b}\cdot\boldsymbol{\eta} + \mathbb{I}[\boldsymbol{\eta}\geq\mathbf{0}]\Big] \\ &= \sup_{\boldsymbol{\mu}\in\mathbb{R}^K}\Big[-R_\oplus(\boldsymbol{\mu}) + \boldsymbol{q}\cdot\boldsymbol{\mu} - \mathbb{I}[\mathbf{A}^\top\boldsymbol{\mu}-\boldsymbol{b}\geq\mathbf{0}]\Big] \\ &= \sup_{\boldsymbol{\mu}\in\mathbb{R}^K}\Big[\boldsymbol{q}\cdot\boldsymbol{\mu} - \Big(R_\oplus(\boldsymbol{\mu}) + \mathbb{I}[\mathbf{A}^\top\boldsymbol{\mu}\geq\boldsymbol{b}]\Big)\Big] \ ,\end{aligned}$$

showing that

$$R_\oplus(\boldsymbol{\mu}) + \mathbb{I}[\mathbf{A}^\top\boldsymbol{\mu}\geq\boldsymbol{b}]$$

is the conjugate of C and the infimum in $\boldsymbol{\eta}$ is attained. To finish the proof we need to verify that the conditions of Theorem 5 hold.

Note that f and g are closed and convex and g is polyhedral. Therefore it remains to show that there exists $\boldsymbol{\mu}\in\operatorname{relint}(\operatorname{dom} f^*)$ such that $\mathbf{A}^\top\boldsymbol{\mu}\in\operatorname{dom} g^*$. Since $\operatorname{dom} f^* = \operatorname{dom} R_\oplus$ and $\mathbf{A}^\top\boldsymbol{\mu}\in\operatorname{dom} g^*$ if and only if $\boldsymbol{\mu}\in\mathcal{M}$, it suffices to show that $\operatorname{relint}(\operatorname{dom} R_\oplus)\cap\mathcal{M}\neq\emptyset$.

Let $\mathcal{M}_g := \{\boldsymbol{\mu}_g : \boldsymbol{\mu}\in\mathcal{M}\}$ and $\mathcal{M}_\oplus := \prod_{g\in\mathcal{G}}\mathcal{M}_g$. By assumption, costs C_g are arbitrage-free, i.e., $\operatorname{dom} R_g = \mathcal{M}_g$. For each g, pick $\tilde{\boldsymbol{\mu}}_g\in\operatorname{relint}\mathcal{M}_g$. Since $\mathcal{M}_g$ is the projection of $\mathcal{M}$ on the coordinate block g, there must exist $\boldsymbol{\mu}^{(g)}\in\mathcal{M}$ such that $\boldsymbol{\mu}^{(g)}_g = \tilde{\boldsymbol{\mu}}_g$. Now, let

$$\boldsymbol{\mu}^\star = \frac{1}{|\mathcal{G}|}\sum_{g\in\mathcal{G}}\boldsymbol{\mu}^{(g)} \ .$$

Note that for $g'\neq g$, we have $\boldsymbol{\mu}^{(g')}_g\in\mathcal{M}_g$, whereas $\boldsymbol{\mu}^{(g)}_g\in\operatorname{relint}\mathcal{M}_g$, so $\boldsymbol{\mu}^\star_g\in\operatorname{relint}\mathcal{M}_g$ and hence $\boldsymbol{\mu}^\star\in\operatorname{relint}\mathcal{M}_\oplus = \operatorname{relint}(\operatorname{dom} R_\oplus)$. At the same time $\boldsymbol{\mu}^\star\in\mathcal{M}$, showing that $\operatorname{relint}(\operatorname{dom} R_\oplus)\cap\mathcal{M}\neq\emptyset$.

Part (c) Fix $\boldsymbol{q}$. Let $\boldsymbol{\eta}^\star$ be a minimizer of Eq. (17) and let $\boldsymbol{\delta}^\star = \mathbf{A}\boldsymbol{\eta}^\star$. Using Theorem 6a, we obtain

$$\begin{aligned}&D(\boldsymbol{\mu}\|\boldsymbol{q}) \\ &\quad= R(\boldsymbol{\mu}) + C(\boldsymbol{q}) - \boldsymbol{\mu}\cdot\boldsymbol{q} \\ &\quad= \mathbb{I}\left[\boldsymbol{\mu}\in\mathcal{M}\right] + R_\oplus(\boldsymbol{\mu}) + C_\oplus(\boldsymbol{q}+\boldsymbol{\delta}^\star) \\ &\qquad\quad - \boldsymbol{b}\cdot\boldsymbol{\eta}^\star - \boldsymbol{\mu}\cdot\boldsymbol{q} \\ &\quad= \mathbb{I}\left[\boldsymbol{\mu}\in\mathcal{M}\right] + R_\oplus(\boldsymbol{\mu}) + C_\oplus(\boldsymbol{q}+\boldsymbol{\delta}^\star) \\ &\qquad\quad - \boldsymbol{\mu}\cdot(\boldsymbol{q}+\boldsymbol{\delta}^\star) + \boldsymbol{\mu}\cdot\boldsymbol{\delta}^\star - \boldsymbol{b}\cdot\boldsymbol{\eta}^\star \\ &\quad= \mathbb{I}\left[\boldsymbol{\mu}\in\mathcal{M}\right] + D_\oplus(\boldsymbol{\mu}\|\boldsymbol{q}+\boldsymbol{\delta}^\star) + (\mathbf{A}^\top\boldsymbol{\mu}-\boldsymbol{b})\cdot\boldsymbol{\eta}^\star \ .\end{aligned}$$

Part (d) If $\boldsymbol{\eta}$ is a minimizer of Eq. (17) then choosing $\boldsymbol{\mu}\in\nabla C(\boldsymbol{q})$, we have $D(\boldsymbol{\mu}\|\boldsymbol{q}) = 0$ and hence by Theorem 6c

$$0 = D(\boldsymbol{\mu}\|\boldsymbol{q}) = D_\oplus(\boldsymbol{\mu}\|\boldsymbol{q}+\boldsymbol{\delta}^\star) + (\mathbf{A}^\top\boldsymbol{\mu}-\boldsymbol{b})\cdot\boldsymbol{\eta}^\star \tag{18}$$

because, by Theorem 6a, C is arbitrage-free, so $\boldsymbol{\mu}\in\mathcal{M}$.

For a converse, assume that for some $\boldsymbol{\mu} \in \mathcal{M}$, $\boldsymbol{\eta} \geq \mathbf{0}$, we have:

$$\begin{aligned}
0 &= D_{\oplus}(\boldsymbol{\mu} \| \boldsymbol{q} + \mathbf{A}\boldsymbol{\eta}) + (\mathbf{A}^{\top}\boldsymbol{\mu} - \boldsymbol{b}) \cdot \boldsymbol{\eta} \\
&= R_{\oplus}(\boldsymbol{\mu}) + C_{\oplus}(\boldsymbol{q} + \mathbf{A}\boldsymbol{\eta}) - \boldsymbol{\mu} \cdot (\boldsymbol{q} + \mathbf{A}\boldsymbol{\eta}) \\
&\qquad + (\mathbf{A}^{\top}\boldsymbol{\mu} - \boldsymbol{b}) \cdot \boldsymbol{\eta} \\
&= R(\boldsymbol{\mu}) + C_{\oplus}(\boldsymbol{q} + \mathbf{A}\boldsymbol{\eta}) - \boldsymbol{\mu} \cdot \boldsymbol{q} - \boldsymbol{b} \cdot \boldsymbol{\eta} \\
&\geq R(\boldsymbol{\mu}) + C(\boldsymbol{q}) - \boldsymbol{\mu} \cdot \boldsymbol{q} \qquad (19) \\
&= D(\boldsymbol{\mu} \| \boldsymbol{q}) \ ,
\end{aligned}$$

where Eq. (19) is from the definition of C. However, since $D(\boldsymbol{\mu} \| \boldsymbol{q}) \geq 0$, we have that Eq. (19) holds with the equality and hence $\boldsymbol{\eta}$ is indeed the minimizer of Eq. (17). ☐

D.2 PROOF OF THEOREM 3

In what follows, let $C_g^t(\boldsymbol{q}_g) = \beta_g(t) C_g(\boldsymbol{q}_g / \beta_g(t))$ and let R_g^t and D_g^t denote the conjugate and divergence derived from C_g^t. Define $\tilde{C}_g$, $\tilde{R}_g$, and $\tilde{D}_g$ similarly.

The definitions of C^t and $\tilde{C}$ imply that

$$\tilde{C}_g(\boldsymbol{q}_g) = \alpha_g C_g^t(\boldsymbol{q}_g / \alpha_g) \ , \qquad (20)$$

$$\tilde{R}_g(\boldsymbol{\mu}_g) = \alpha_g R_g^t(\boldsymbol{\mu}_g) \ , \qquad (21)$$

$$\tilde{D}_g(\boldsymbol{\mu}_g \| \boldsymbol{q}_g) = \alpha_g D_g^t(\boldsymbol{\mu}_g \| \boldsymbol{q}_g / \alpha_g) \ . \qquad (22)$$

The proof proceeds in several steps:

Step 1 *$D^t(\boldsymbol{\mu} \| \boldsymbol{s}) = 0$ if and only if $\boldsymbol{\mu} \in \mathcal{M}$,*

$$\begin{aligned}
&D_g^t(\boldsymbol{\mu}_g \| \boldsymbol{s}_g + \boldsymbol{\delta}_g^{\star}) = 0 \textit{ for all } g \in \mathcal{G}, \\
&\textit{and } (\mathbf{A}^{\top}\boldsymbol{\mu} - \boldsymbol{b}) \cdot \boldsymbol{\eta}^{\star} = 0.
\end{aligned} \qquad (23)$$

If Eq. (23) holds and $\boldsymbol{\mu} \in \mathcal{M}$, then Theorem 6c shows that $D^t(\boldsymbol{\mu} \| \boldsymbol{s}) = 0$. For the opposite implication note that $D^t(\boldsymbol{\mu} \| \boldsymbol{s}) = \infty$ if $\boldsymbol{\mu} \notin \mathcal{M}$, so we must have $\boldsymbol{\mu} \in \mathcal{M}$. For $\boldsymbol{\mu} \in \mathcal{M}$, by Theorem 6c,

$$D^t(\boldsymbol{\mu} \| \boldsymbol{s}) = \sum_g D_g^t(\boldsymbol{\mu}_g \| \boldsymbol{s}_g + \boldsymbol{\delta}_g^{\star}) + (\mathbf{A}^{\top}\boldsymbol{\mu} - \boldsymbol{b}) \cdot \boldsymbol{\eta}^{\star} \ . \qquad (24)$$

Note that the last term in Eq. (24) is non-negative, because $\boldsymbol{\eta}^{\star} \geq 0$ and $\boldsymbol{\mu} \in \mathcal{M}$. Since also the divergences D_g^t are non-negative, we obtain that all the terms must equal zero if $D^t(\boldsymbol{\mu} \| \boldsymbol{s}) = 0$.

Step 2 $\boldsymbol{\eta}^{\star} \in \operatorname{argmin}_{\boldsymbol{\eta} \geq \mathbf{0}} \left[\tilde{C}_{\oplus}(\tilde{\boldsymbol{s}} + \mathbf{A}\boldsymbol{\eta}) - \boldsymbol{b} \cdot \boldsymbol{\eta} \right]$.

By Theorem 6d, it suffices to exhibit $\boldsymbol{\mu} \in \mathcal{M}$ such that

$$\tilde{D}_{\oplus}(\boldsymbol{\mu} \| \tilde{\boldsymbol{s}} + \boldsymbol{\delta}^{\star}) + (\mathbf{A}^{\top}\boldsymbol{\mu} - \boldsymbol{b}) \cdot \boldsymbol{\eta}^{\star} = 0 \ . \qquad (25)$$

Pick any $\boldsymbol{\mu} \in \partial C^t(\boldsymbol{s})$, i.e., $D^t(\boldsymbol{\mu} \| \boldsymbol{s}) = 0$. Then by expanding $\tilde{D}_{\oplus}$ (using Eq. 22) and then using the definition of $\tilde{\boldsymbol{s}}$, we obtain

$$\begin{aligned}
&\tilde{D}_{\oplus}(\boldsymbol{\mu} \| \tilde{\boldsymbol{s}} + \boldsymbol{\delta}^{\star}) + (\mathbf{A}^{\top}\boldsymbol{\mu} - \boldsymbol{b}) \cdot \boldsymbol{\eta}^{\star} \\
&\quad = \sum_g \alpha_g D_g^t\Big(\boldsymbol{\mu}_g \,\Big\|\, \frac{\tilde{\boldsymbol{s}}_g + \boldsymbol{\delta}_g^{\star}}{\alpha_g}\Big) + (\mathbf{A}^{\top}\boldsymbol{\mu} - \boldsymbol{b}) \cdot \boldsymbol{\eta}^{\star} \\
&\quad = \sum_g \alpha_g D_g^t(\boldsymbol{\mu}_g \| \boldsymbol{s}_g + \boldsymbol{\delta}_g^{\star}) + (\mathbf{A}^{\top}\boldsymbol{\mu} - \boldsymbol{b}) \cdot \boldsymbol{\eta}^{\star}.
\end{aligned}$$

Both terms on the right-hand side are zero by Step 1, yielding Eq. (25) as desired.

Step 3 *For all $\boldsymbol{\mu} \in \mathcal{M}$:*

$$\tilde{D}(\boldsymbol{\mu} \| \tilde{\boldsymbol{s}}) = \textstyle\sum_g \alpha_g D_g^t(\boldsymbol{\mu}_g \| \boldsymbol{s}_g + \boldsymbol{\delta}_g^{\star}) + (\mathbf{A}^{\top}\boldsymbol{\mu} - \boldsymbol{b}) \cdot \boldsymbol{\eta}^{\star} \ .$$

This follows by Step 2 and Theorem 6c plus Eq. 22, noting that

$$(\tilde{\boldsymbol{s}}_g + \boldsymbol{\delta}_g^{\star}) / \alpha_g = \boldsymbol{s}_g + \boldsymbol{\delta}_g^{\star} \ .$$

Step 4 *$\tilde{C}$ and $\tilde{\boldsymbol{s}}$ satisfy* PRICE.

Since $\boldsymbol{\mu} \in \partial C^t(\boldsymbol{s})$ if and only if $D^t(\boldsymbol{\mu} \| \boldsymbol{s}) = 0$, and similarly for $\boldsymbol{\mu} \in \partial \tilde{C}(\tilde{\boldsymbol{s}})$, it suffices to show that $D^t(\boldsymbol{\mu} \| \boldsymbol{s}) = 0$ if and only if $\tilde{D}(\boldsymbol{\mu} \| \tilde{\boldsymbol{s}}) = 0$. First assume that $D^t(\boldsymbol{\mu} \| \boldsymbol{s}) = 0$. Then Steps 1 and 3 show that $\tilde{D}(\boldsymbol{\mu} \| \tilde{\boldsymbol{s}}) = 0$. Also, vice versa: if $\tilde{D}(\boldsymbol{\mu} \| \tilde{\boldsymbol{s}}) = 0$ then, from Step 3 (by a similar reasoning as in the proof of Step 1), we have that $\boldsymbol{\mu} \in \mathcal{M}$, for all g it holds that $D_g^t(\boldsymbol{\mu}_g \| \boldsymbol{s}_g + \boldsymbol{\delta}_g^{\star}) = 0$, and also $(\mathbf{A}^{\top}\boldsymbol{\mu} - \boldsymbol{b}) \cdot \boldsymbol{\eta}^{\star} = 0$. Hence, by Step 1, also $D^t(\boldsymbol{\mu} \| \boldsymbol{s}) = 0$.

D.3 PROOF OF THEOREM 4

We first show that CONDPRICE and EXUTIL hold. We proceed by Prop. 1. Fix $\boldsymbol{x} \in \mathcal{X}_g$, let $\Omega^{\boldsymbol{x}} = \{\boldsymbol{\rho}_g = \boldsymbol{x}\}$, and let $\boldsymbol{\mu} \in \mathcal{M}^{\boldsymbol{x}} := \mathcal{M}(\Omega^{\boldsymbol{x}})$. Then, expanding $D^t(\boldsymbol{\mu} \| \boldsymbol{s})$ according to Theorem 6c and $\tilde{D}(\boldsymbol{\mu} \| \tilde{\boldsymbol{s}})$ according to Theorem 3, we have

$$\begin{aligned}
D^t(\boldsymbol{\mu} \| \boldsymbol{s}) - \tilde{D}(\boldsymbol{\mu} \| \tilde{\boldsymbol{s}}) &= (1 - \alpha_g) D_g^t(\boldsymbol{\mu}_g \| \boldsymbol{s}_g + \boldsymbol{\delta}_g^{\star}) \\
&= (1 - \alpha_g) D_g^t(\boldsymbol{x} \| \boldsymbol{s}_g + \boldsymbol{\delta}_g^{\star}) \qquad (26)
\end{aligned}$$

which is a constant independent of the specific choice of $\boldsymbol{\mu} \in \mathcal{M}^{\boldsymbol{x}}$, proving that both CONDPRICE and EXUTIL hold.

Next, we show that DECUTIL holds. Let $\hat{\boldsymbol{\mu}}^{\boldsymbol{x}} \in \boldsymbol{p}^t(\Omega^{\boldsymbol{x}}; \boldsymbol{s}) = \tilde{\boldsymbol{p}}(\Omega^{\boldsymbol{x}}; \tilde{\boldsymbol{s}})$ (the equality holds by EXUTIL). From Eq. (26) and Theorem 1, we have

$$\begin{aligned}
&\tilde{\mathtt{Util}}(\boldsymbol{\rho}_g = \boldsymbol{x}; \tilde{\boldsymbol{s}}) = \tilde{D}(\hat{\boldsymbol{\mu}}^{\boldsymbol{x}} \| \tilde{\boldsymbol{s}}) \\
&\quad = D^t(\hat{\boldsymbol{\mu}}^{\boldsymbol{x}} \| \boldsymbol{s}) - (1 - \alpha_g) D_g^t(\boldsymbol{x} \| \boldsymbol{s}_g + \boldsymbol{\delta}_g^{\star}) \\
&\quad = \mathtt{Util}(\boldsymbol{\rho}_g = \boldsymbol{x}; \boldsymbol{s}) - (1 - \alpha_g) D_g^t(\boldsymbol{x} \| \boldsymbol{s}_g + \boldsymbol{\delta}_g^{\star}) \ ,
\end{aligned}$$

i.e., the utility for event $\Omega^{\boldsymbol{x}}$ is non-increasing, because $\alpha_g \in (0, 1)$.

In order to show DECUTIL, we still need to show that $D_g^t(\boldsymbol{x}\|\boldsymbol{s}_g + \boldsymbol{\delta}_g^\star) = 0$ implies $D^t(\hat{\boldsymbol{\mu}}^{\boldsymbol{x}}\|\boldsymbol{s}) = 0$. Assume that C_g^t is differentiable and the submarket g is tight, i.e., $\mathcal{M}^{\boldsymbol{x}} = \{\boldsymbol{\mu} \in \mathcal{M} : \boldsymbol{\mu}_g = \boldsymbol{x}\}$. Assume that $D_g^t(\boldsymbol{x}\|\boldsymbol{s}_g + \boldsymbol{\delta}_g^\star) = 0$. By differentiability of C_g^t this implies that $\boldsymbol{x} = \nabla C_g^t(\boldsymbol{s}_g + \boldsymbol{\delta}_g^\star)$, and hence $\boldsymbol{\mu}_g = \boldsymbol{x}$ for all $\boldsymbol{\mu} \in \partial C^t(\boldsymbol{s})$. By assumption, any of them is in $\mathcal{M}^{\boldsymbol{x}}$ and hence any of them can be chosen as a minimizer $\hat{\boldsymbol{\mu}}^{\boldsymbol{x}}$ with $D^t(\hat{\boldsymbol{\mu}}^{\boldsymbol{x}}\|\boldsymbol{s}) = 0$.

D.4 BINARY-PAYOFF SUBMARKETS ARE TIGHT

Theorem 7. *Let g be a binary-payoff submarket in an LCMM, i.e., $\boldsymbol{\rho}_g(\omega) \in \{0,1\}^g$ for all $\omega \in \Omega$. Then g is tight.*

Proof. Fix any $\boldsymbol{x} \in \mathcal{X}_g = \{0,1\}^g$. Let $\Omega^{\boldsymbol{x}} \coloneqq \{\boldsymbol{\rho}_g = \boldsymbol{x}\}$, and let $\mathcal{M}^{\boldsymbol{x}} \coloneqq \mathcal{M}(\Omega^{\boldsymbol{x}})$ be the set of beliefs consistent with $\Omega^{\boldsymbol{x}}$. Let $\boldsymbol{\mu} \in \mathcal{M}$ be such that $\boldsymbol{\mu}_g = \boldsymbol{x}$. We need to show that $\boldsymbol{\mu} \in \mathcal{M}^{\boldsymbol{x}}$.

Since $\boldsymbol{\mu} \in \mathcal{M}$, we can write $\boldsymbol{\mu} = \sum_{\omega\in\Omega} \lambda_\omega \boldsymbol{\rho}(\omega)$ for some $\lambda_\omega \ge 0$ such that $\sum_{\omega\in\Omega} \lambda_\omega = 1$. We will argue that the condition $\boldsymbol{\mu}_g = \boldsymbol{x}$ implies that $\lambda_\omega = 0$ for $\omega \notin \Omega^{\boldsymbol{x}}$ and thus in fact $\boldsymbol{\mu} \in \mathcal{M}^{\boldsymbol{x}}$. Essentially we show that $\mathcal{M}^{\boldsymbol{x}}$ is the set of maximizers of a linear function over $\mathcal{M}$ and that $\boldsymbol{\mu}$ is one of the maximizers.

The required linear function, $\boldsymbol{v}\cdot\boldsymbol{\mu}$, is specified by the vector $\boldsymbol{v} \in \mathbb{R}^K$ defined as follows:

$$v_i = \begin{cases} 1 & \text{if } i \in g \text{ and } x_i = 1, \\ -1 & \text{if } i \in g \text{ and } x_i = 0, \\ 0 & \text{if } i \notin g. \end{cases}$$

Let k be the number of 1s in the vector $\boldsymbol{x}$. From the definition of $\Omega^{\boldsymbol{x}}$ we have that $\boldsymbol{v}\cdot\boldsymbol{\rho}(\omega) = k$ for all $\omega \in \Omega^{\boldsymbol{x}}$. Let $\omega' \notin \Omega^{\boldsymbol{x}}$, i.e., $\omega' \in \Omega^{\boldsymbol{x}'}$ for some $\boldsymbol{x}' \in \mathcal{X}_g\backslash\{\boldsymbol{x}\}$. Since $\boldsymbol{x}' \in \{0,1\}^g$ but $\boldsymbol{x}' \ne \boldsymbol{x}$, there exists $i \in g$ such that $x_i = 1$ but $x_i' = 0$, or such that $x_i = 0$ but $x_i' = 1$. Thus, $\boldsymbol{v}_g \cdot \boldsymbol{x}' \le k-1$ and hence also $\boldsymbol{v}\cdot\boldsymbol{\rho}(\omega') \le k-1$. This yields

$$\begin{aligned} \boldsymbol{v}\cdot\boldsymbol{\mu} = \boldsymbol{v}\cdot\left(\sum_{\omega\in\Omega}\lambda_\omega\boldsymbol{\rho}(\omega)\right) &= \sum_{\omega\in\Omega}\lambda_\omega\Big(\boldsymbol{v}\cdot\boldsymbol{\rho}(\omega)\Big) \\ &\le k\left(\sum_{\omega\in\Omega^{\boldsymbol{x}}}\lambda_\omega\right) + (k-1)\left(\sum_{\omega'\in\Omega\backslash\Omega^{\boldsymbol{x}}}\lambda_{\omega'}\right) \\ &= k - \sum_{\omega'\in\Omega\backslash\Omega^{\boldsymbol{x}}}\lambda_{\omega'} \enspace . \end{aligned}$$

However, $\boldsymbol{\mu}_g = \boldsymbol{x}$ and thus $\boldsymbol{v}\cdot\boldsymbol{\mu} = k$. Therefore, we must have $\lambda_{\omega'} = 0$ for $\omega' \in \Omega\backslash\Omega^{\boldsymbol{x}}$, proving the theorem. ☐

E CONDITIONAL PRICE VECTORS

E.1 CONDITIONAL PRICES FOR LMSR

In this section we show that conditional price vectors for LMSR coincide with conditional probabilities.

Recall that for LMSR, we have the outcomes $\Omega = [K]$, payoffs $\rho_i(\omega) = \mathbf{1}[\omega = i]$, and prices

$$p_i(\boldsymbol{q}) = \frac{e^{q_i}}{\sum_{j\in[K]} e^{q_j}} \enspace ,$$

i.e., price vectors are probability distributions over $i \in [K]$.

The mixed Bregman divergence for LMSR has the form

$$D(\boldsymbol{\mu}\|\boldsymbol{q}) = \sum_{i\in[K]} \mu_i \ln\left(\frac{\mu_i}{p_i(\boldsymbol{q})}\right) = \mathrm{KL}\big(\boldsymbol{\mu} \,\big\|\, \boldsymbol{p}(\boldsymbol{q})\big)$$

where $\mathrm{KL}(\boldsymbol{\mu}\|\boldsymbol{\nu}) \coloneqq \sum_{i\in[K]} \mu_i \ln(\mu_i/\nu_i)$ is the KL divergence defined for any pair of distributions $\boldsymbol{\mu},\boldsymbol{\nu}$ on $[K]$. KL divergence is always non-negative, possibly equal to ∞, and equal to zero if and only if $\boldsymbol{\mu} = \boldsymbol{\nu}$.

Let $\boldsymbol{q} \in \mathbb{R}^K$ and $\mathcal{E} \subseteq [K] = \Omega$ be a non-null event. Let $\hat{\boldsymbol{\mu}}$ be the probability vector obtained by conditioning $\boldsymbol{p}(\boldsymbol{q})$ on the event $\mathcal{E}$, i.e.,

$$\hat{\mu}_i = \begin{cases} p_i(\boldsymbol{q})/c & \text{if } i \in \mathcal{E}, \\ 0 & \text{otherwise,} \end{cases}$$

where $c = \sum_{i\in\mathcal{E}} p_i(\boldsymbol{q})$ is the normalization over $\mathcal{E}$. Note that $p_i(\boldsymbol{q}) > 0$ for all $i \in [K]$, so $c > 0$. We will now argue that $\boldsymbol{p}(\mathcal{E};\boldsymbol{q}) = \{\hat{\boldsymbol{\mu}}\}$.

We appeal to Eq. (3) of Theorem 1. Specifically, we will show that $\hat{\boldsymbol{\mu}}$ is the unique minimizer of $\min_{\boldsymbol{\mu}'\in\mathcal{M}(\mathcal{E})} D(\boldsymbol{\mu}'\|\boldsymbol{q})$.

First, note that $\hat{\boldsymbol{\mu}} \in \mathcal{M}(\mathcal{E})$, and from the definition of $\hat{\boldsymbol{\mu}}$

$$\begin{aligned} D(\hat{\boldsymbol{\mu}}\|\boldsymbol{q}) = \sum_{i\in\mathcal{E}} \hat{\mu}_i \ln\left(\frac{\hat{\mu}_i}{p_i(\boldsymbol{q})}\right) &= \sum_{i\in\mathcal{E}} \hat{\mu}_i \ln\left(\frac{p_i(\boldsymbol{q})/c}{p_i(\boldsymbol{q})}\right) \\ &= \sum_{i\in\mathcal{E}} \hat{\mu}_i \ln(1/c) = \ln(1/c) \enspace . \end{aligned}$$

Now, let $\boldsymbol{\mu}' \in \mathcal{M}(\mathcal{E})$ and compare the values $D(\boldsymbol{\mu}'\|\boldsymbol{q})$ and $D(\hat{\boldsymbol{\mu}}\|\boldsymbol{q})$:

$$\begin{aligned} D(\boldsymbol{\mu}'\|\boldsymbol{q}) - D(\hat{\boldsymbol{\mu}}\|\boldsymbol{q}) &= \left(\sum_{i\in\mathcal{E}} \mu_i' \ln\left(\frac{\mu_i'}{p_i(\boldsymbol{q})}\right)\right) - \ln(1/c) \\ &= \left(\sum_{i\in\mathcal{E}} \mu_i' \ln\left(\frac{\mu_i'}{p_i(\boldsymbol{q})}\right)\right) - \left(\sum_{i\in\mathcal{E}} \mu_i' \ln(1/c)\right) \\ &= \sum_{i\in\mathcal{E}} \mu_i' \ln\left(\frac{\mu_i'}{p_i(\boldsymbol{q})/c}\right) \\ &= \sum_{i\in\mathcal{E}} \mu_i' \ln\left(\frac{\mu_i'}{\hat{\mu}_i}\right) = \mathrm{KL}(\boldsymbol{\mu}'\|\hat{\boldsymbol{\mu}}) \enspace . \end{aligned}$$

Thus, we have $D(\boldsymbol{\mu}' \| \boldsymbol{q}) \geq D(\hat{\boldsymbol{\mu}} \| \boldsymbol{q})$ with equality if and only if $\boldsymbol{\mu}' = \hat{\boldsymbol{\mu}}$, i.e., $\hat{\boldsymbol{\mu}}$ is the sole minimizer of $\min_{\boldsymbol{\mu}' \in \mathcal{M}(\mathcal{E})} D(\boldsymbol{\mu}' \| \boldsymbol{q})$.

E.2 OPTIMAL TRADING GIVEN $\mathcal{E}$

In this section we analyze the prices that result from actions of a trader optimizing his guaranteed profit from the information $\omega \in \mathcal{E}$ as in Definition 2 in the market with cost function C. Intuitively, we would like to say that such a trader would move the market price to a conditional price vector $\hat{\boldsymbol{\mu}} \in \boldsymbol{p}(\mathcal{E}; \boldsymbol{q})$. However, this may not be possible. For example, consider a complete market using LMSR. In such a market, a trader can push the market price arbitrarily close to any $\boldsymbol{\mu} \in \mathcal{M}(\mathcal{E})$, but cannot push the price all the way to $\boldsymbol{\mu}$ with any finite purchase (unless $\mathcal{E} = \Omega$).

Because of this, instead of reasoning directly about finite purchases, we introduce the notion of an *optimizing action sequence* and show that in the limit such a trader would move the market from a state $\boldsymbol{q}$ to states that minimize the Bregman divergence to conditional price vectors $\hat{\boldsymbol{\mu}} \in \boldsymbol{p}(\mathcal{E}; \boldsymbol{q})$. Then we argue that for R strictly convex (such as entropy in case of LMSR), this implies that the resulting market price vector approaches the unique conditional price vector in the limit.

We begin by formalizing the optimizing behavior in Definition 2.

Definition 8. *We say that $\{\boldsymbol{r}_i\}_{i=1}^{\infty}$ is an* optimizing action sequence *with respect to a non-null event $\mathcal{E}$ and a state $\boldsymbol{q}$ if*

$$\lim_{i \to \infty} \min_{\omega \in \mathcal{E}} \Big[\boldsymbol{\rho}(\omega) \cdot \boldsymbol{r}_i - C(\boldsymbol{q} + \boldsymbol{r}_i) + C(\boldsymbol{q}) \Big] = \texttt{Util}(\mathcal{E}; \boldsymbol{q}) \ .$$

We say that $\{\boldsymbol{q}_i\}_{i=1}^{\infty}$ is an optimizing state sequence *with respect to $\mathcal{E}$ and $\boldsymbol{q}$ if*

$$\lim_{i \to \infty} \min_{\omega \in \mathcal{E}} \Big[\boldsymbol{\rho}(\omega) \cdot (\boldsymbol{q}_i - \boldsymbol{q}) - C(\boldsymbol{q}_i) + C(\boldsymbol{q}) \Big] = \texttt{Util}(\mathcal{E}; \boldsymbol{q}) \ .$$

(Thus, any optimizing action sequence yields an optimizing state sequence $\boldsymbol{q}_i = \boldsymbol{q} + \boldsymbol{r}_i$ and vice versa.)

We next show that optimizing state sequences minimize divergence to conditional price vectors. Specifically, the divergence between any state sequence and any conditional price vector tends to zero. Loosely speaking, this means that the market is moving towards states whose associated prices, in the limit, include all conditional price vectors.

Theorem 8. *Let $\{\boldsymbol{q}_i\}_{i=1}^{\infty}$ be an optimizing state sequence with respect to $\mathcal{E}$ and $\boldsymbol{q}$, and let $\hat{\boldsymbol{\mu}} \in \boldsymbol{p}(\mathcal{E}; \boldsymbol{q})$. Then $D(\hat{\boldsymbol{\mu}} \| \boldsymbol{q}_i) \to 0$ as $i \to \infty$.*

Proof. Since the minimized objective in Definition 8 is linear in $\boldsymbol{\rho}(\omega)$, we can without loss of generality replace minimization over $\boldsymbol{\rho}(\omega)$ where $\omega \in \mathcal{E}$ by minimization over $\boldsymbol{\mu}' \in \mathcal{M}(\mathcal{E})$, and thus assume that

$$\lim_{i \to \infty} \min_{\boldsymbol{\mu}' \in \mathcal{M}(\mathcal{E})} \Big[\boldsymbol{\mu}' \cdot (\boldsymbol{q}_i - \boldsymbol{q}) - C(\boldsymbol{q}_i) + C(\boldsymbol{q}) \Big] = \texttt{Util}(\mathcal{E}; \boldsymbol{q}) \,. \tag{27}$$

The expression in the brackets can be rewritten as

$$\boldsymbol{\mu}' \cdot (\boldsymbol{q}_i - \boldsymbol{q}) - C(\boldsymbol{q}_i) + C(\boldsymbol{q}) = -D(\boldsymbol{\mu}' \| \boldsymbol{q}_i) + D(\boldsymbol{\mu}' \| \boldsymbol{q}) \,.$$

Furthermore, by Theorem 1, we have $\texttt{Util}(\mathcal{E}; \boldsymbol{q}) = D(\hat{\boldsymbol{\mu}} \| \boldsymbol{q})$. We can therefore rewrite Eq. (27) as

$$\lim_{i \to \infty} \min_{\boldsymbol{\mu}' \in \mathcal{M}(\mathcal{E})} \Big[D(\boldsymbol{\mu}' \| \boldsymbol{q}) - D(\boldsymbol{\mu}' \| \boldsymbol{q}_i) \Big] = D(\hat{\boldsymbol{\mu}} \| \boldsymbol{q}) \,. \tag{28}$$

To get the statement of the theorem, note that for all i,

$$D(\hat{\boldsymbol{\mu}} \| \boldsymbol{q}) \geq D(\hat{\boldsymbol{\mu}} \| \boldsymbol{q}) - D(\hat{\boldsymbol{\mu}} \| \boldsymbol{q}_i) \tag{29}$$

$$\geq \min_{\boldsymbol{\mu}' \in \mathcal{M}(\mathcal{E})} \Big[D(\boldsymbol{\mu}' \| \boldsymbol{q}) - D(\boldsymbol{\mu}' \| \boldsymbol{q}_i) \Big] \tag{30}$$

where Eq. (29) follows by non-negativity of the divergence, and Eq. (30) because $\hat{\boldsymbol{\mu}} \in \mathcal{M}(\mathcal{E})$. Since Eq. (30) converges to $D(\hat{\boldsymbol{\mu}} \| \boldsymbol{q})$ by Eq. (28), we obtain that the right hand-side in Eq. (29) must also converge to $D(\hat{\boldsymbol{\mu}} \| \boldsymbol{q})$, i.e., $D(\hat{\boldsymbol{\mu}} \| \boldsymbol{q}_i) \to 0$. □

When R is strictly convex on $\mathcal{M}$, Theorem 8 can be strengthened to show that $\boldsymbol{p}(\boldsymbol{q}_i) \to \hat{\boldsymbol{\mu}}$. Strict convexity of R is equivalent to a certain notion of smoothness of C. It is stronger than differentiability of C [27, Theorem 26.3], but weaker than the existence of a Lipschitz-continuous gradient for C. [4]

Theorem 9. *Let $\{\boldsymbol{q}_i\}_{i=1}^{\infty}$ be an optimizing state sequence with respect to $\mathcal{E}$ and $\boldsymbol{q}$, and let $\hat{\boldsymbol{\mu}} \in \boldsymbol{p}(\mathcal{E}; \boldsymbol{q})$. If R is strictly convex on $\mathcal{M}$ then $\boldsymbol{p}(\boldsymbol{q}_i) \to \hat{\boldsymbol{\mu}}$ as $i \to \infty$.*

Proof. First note that if R is strictly convex then C is differentiable [27, Theorem 26.3], and thus $\boldsymbol{p}(\boldsymbol{q}_i)$ is always a singleton. Next note that the sequence $\{\boldsymbol{p}(\boldsymbol{q}_i)\}_{i=1}^{\infty}$ is contained in a compact set $\mathcal{M}$, so it must have a cluster point in $\mathcal{M}$. Pick an arbitrary cluster point $\boldsymbol{\mu}^\star$ and choose a subsequence $\{\boldsymbol{q}_{i(j)}\}_{j=1}^{\infty}$ such that $\boldsymbol{p}(\boldsymbol{q}_{i(j)}) \to \boldsymbol{\mu}^\star$ as $j \to \infty$. We will show that $\boldsymbol{\mu}^\star = \hat{\boldsymbol{\mu}}$ and thus all of the cluster points of the original price sequence $\{\boldsymbol{p}(\boldsymbol{q}_i)\}_{i=1}^{\infty}$ coincide. This implies that the sequence actually converges to $\hat{\boldsymbol{\mu}}$ (again, because it is contained in a compact set $\mathcal{M}$).

To simplify writing, let $\boldsymbol{q}'_j := \boldsymbol{q}_{i(j)}$ and $\boldsymbol{\mu}'_j := \boldsymbol{p}(\boldsymbol{q}'_j)$. By the choice of the subsequence, we have $\boldsymbol{\mu}'_j \to \boldsymbol{\mu}^\star$. By Prop. 5, we have $\boldsymbol{q}'_j \in \partial R(\boldsymbol{\mu}'_j)$ and by convexity of R we have the lower bound

$$R(\boldsymbol{\mu}) \geq R(\boldsymbol{\mu}'_j) + (\boldsymbol{\mu} - \boldsymbol{\mu}'_j) \cdot \boldsymbol{q}'_j \tag{31}$$

[4] Proposition 12.60ab, R. Tyrrell Rockafellar, Roger J.-B. Wets. *Variational analysis*. Springer, 1998.

valid for all $\boldsymbol{\mu}$. We will analyze the limits of this lower bound on the line segment connecting $\boldsymbol{\mu}^\star$ and $\hat{\boldsymbol{\mu}}$ to argue that R must be linear on this line segment. This will yield a contradiction unless $\boldsymbol{\mu}^\star = \hat{\boldsymbol{\mu}}$.

By Theorem 8 we have that $D(\hat{\boldsymbol{\mu}} \| \boldsymbol{q}_i) \to 0$ and hence also $D(\hat{\boldsymbol{\mu}} \| \boldsymbol{q}'_j) \to 0$. To begin the analysis of the lower bound in Eq. (31), we rewrite $D(\hat{\boldsymbol{\mu}} \| \boldsymbol{q}'_j)$ as

$$\begin{aligned} D(\hat{\boldsymbol{\mu}} \| \boldsymbol{q}'_j) &= R(\hat{\boldsymbol{\mu}}) + C(\boldsymbol{q}'_j) - \hat{\boldsymbol{\mu}} \cdot \boldsymbol{q}'_j \\ &= R(\hat{\boldsymbol{\mu}}) - R(\boldsymbol{\mu}'_j) + (\boldsymbol{\mu}'_j - \hat{\boldsymbol{\mu}}) \cdot \boldsymbol{q}'_j \end{aligned} \tag{32}$$

where the last equality follows because $C(\boldsymbol{q}'_j) = \boldsymbol{\mu}'_j \cdot \boldsymbol{q}'_j - R(\boldsymbol{\mu}'_j)$ by Prop. 5. Since $D(\hat{\boldsymbol{\mu}} \| \boldsymbol{q}'_j) \to 0$, Eq. (32) yields

$$\lim_{j \to \infty} \left[R(\boldsymbol{\mu}'_j) + (\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}'_j) \cdot \boldsymbol{q}'_j \right] = R(\hat{\boldsymbol{\mu}}) \ . \tag{33}$$

Thus, we see that the lower bound of Eq. (31) at $\boldsymbol{\mu} = \hat{\boldsymbol{\mu}}$ is tight as $j \to \infty$.

Next, we note that R is continuous on $\mathcal{M}$ by Prop. 3, because $\mathcal{M}$ is polyhedral.

We now focus on the line segment connecting $\boldsymbol{\mu}^\star$ and $\hat{\boldsymbol{\mu}}$. Let $\lambda \in [0,1]$ and consider Eq. (31) at $\boldsymbol{\mu}'_j(\lambda) := (1 - \lambda)\boldsymbol{\mu}'_j + \lambda \hat{\boldsymbol{\mu}}$:

$$\begin{aligned} R(\boldsymbol{\mu}'_j(\lambda)) &\geq R(\boldsymbol{\mu}'_j) + (\boldsymbol{\mu}'_j(\lambda) - \boldsymbol{\mu}'_j) \cdot \boldsymbol{q}'_j \\ &= R(\boldsymbol{\mu}'_j) + \lambda(\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}'_j) \cdot \boldsymbol{q}'_j \\ &= (1-\lambda) R(\boldsymbol{\mu}'_j) \\ &\qquad + \lambda \Big(R(\boldsymbol{\mu}'_j) + (\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}'_j) \cdot \boldsymbol{q}'_j \Big) \ , \end{aligned}$$

where the first equality follows from the definition of $\boldsymbol{\mu}'_j(\lambda)$ and the second by rearranging the terms. Taking $j \to \infty$ and using Eq. (33) and the continuity of R, we obtain

$$R\Big((1-\lambda)\boldsymbol{\mu}^\star + \lambda \hat{\boldsymbol{\mu}} \Big) \geq (1-\lambda) R(\boldsymbol{\mu}^\star) + \lambda R(\hat{\boldsymbol{\mu}}) \ .$$

However, by convexity we also have

$$R\Big((1-\lambda)\boldsymbol{\mu}^\star + \lambda \hat{\boldsymbol{\mu}} \Big) \leq (1-\lambda) R(\boldsymbol{\mu}^\star) + \lambda R(\hat{\boldsymbol{\mu}}) \ ,$$

so indeed R must be linear on the line segment connecting $\boldsymbol{\mu}^\star$ and $\hat{\boldsymbol{\mu}}$, which contradicts strict convexity of R unless $\boldsymbol{\mu}^\star = \hat{\boldsymbol{\mu}}$. □

F ROBUST BAYES UTILITY

In Sec. 2, we motivate the utility for information as the market maker's willingness to pay for information, or, equivalently, as the traders' ability to profit from their information. Another motivation for the same definitions, pursued in Sec. 2.5, arises from defining the utility for information via a measure of distance, such that the market maker is willing to pay more for the information more distant from the current state.

In this section, we give a fourth motivation, showing how our definitions naturally match up with concepts from robust Bayes decision theory [16]. In Sec. 2, we adopted the perspective of either an expected-utility-maximizing trader (for the utility of a belief) or a worst-case trader (for the utility of an event). Here we show that if we make a slightly stronger assumption about the behavior of traders endowed with various information relevant to the market maker, these two notions can be unified. Specifically, we will show that assuming that the traders are robust Bayes decision makers, we obtain the same definitions of the utility for information.

As before, let Ω be a finite set of outcomes. Let Δ be the set of probability distributions over Ω. Consider a decision maker trying to choose an action a from some action set before an outcome is realized. Given an action a and a realized outcome $\omega \in \Omega$, the decision maker receives the utility $u(a, \omega)$. We assume that the decision maker's information $\mathcal{I}$ is represented as a non-null subset of Δ, i.e., $\emptyset \neq \mathcal{I} \subseteq \Delta$. The decision maker assumes that the outcome ω is drawn according to some probability distribution P, but the only information about P is that $P \in \mathcal{I}$. Given this information, we call the decision maker the *robust Bayes decision maker* if he is trying to maximize the worst-case expected utility where the worst case is over $P \in \mathcal{I}$. The obtained worst-case expected utility is referred to as the *robust Bayes utility for* $\mathcal{I}$ and defined as

$$\texttt{RBUtil}(\mathcal{I}) := \sup_a \inf_{P \in \mathcal{I}} \mathbb{E}_{\omega \sim P}[u(a, \omega)] \ .$$

Consider a prediction market with the cost function C and the current state $\boldsymbol{q}$. Actions available to a trader are all possible trades $\boldsymbol{r} \in \mathbb{R}^K$, and the utility of the trader is

$$u(\boldsymbol{r}, \omega) = \boldsymbol{\rho}(\omega) \cdot \boldsymbol{r} - C(\boldsymbol{q} + \boldsymbol{r}) + C(\boldsymbol{q}) \ .$$

To see that our utility for information is actually the robust Bayes utility, define the following information sets:

$$\begin{aligned} \{\mathbb{E}_P[\boldsymbol{\rho}] = \boldsymbol{\mu}\} &:= \{P \in \Delta : \ \mathbb{E}_P[\boldsymbol{\rho}] = \boldsymbol{\mu}\} \\ \{P[\mathcal{E}] = 1\} &:= \{P \in \Delta : \ P[\mathcal{E}] = 1\} \ . \end{aligned}$$

The first corresponds to the probability distributions P that give rise to the expected value $\mathbb{E}_P[\boldsymbol{\rho}] = \boldsymbol{\mu}$; the second corresponds to the probability distributions that put all of their mass on outcomes $\omega \in \mathcal{E}$. Plugging these information sets into the definition of the robust Bayes utility, we obtain

$$\begin{aligned} \texttt{RBUtil}(\mathbb{E}_P[\boldsymbol{\rho}] = \boldsymbol{\mu}) &= \texttt{Util}(\boldsymbol{\mu}; \boldsymbol{q}) \\ \texttt{RBUtil}(P[\mathcal{E}] = 1) &= \texttt{Util}(\mathcal{E}; \boldsymbol{q}) \ . \end{aligned}$$

Thus indeed the market maker's utility for a belief and for an event is a robust Bayes utility.

While the notion of excess utility is not standard in robust Bayes decision theory, it can be naturally defined as follows. Let $\mathcal{I}_1, \mathcal{I}_2 \subseteq \Omega$ such that $\mathcal{I}_1 \cap \mathcal{I}_2 \neq \emptyset$. Then the *excess robust Bayes utility for* $\mathcal{I}_1$ *given* $\mathcal{I}_2$ is

$$\mathtt{RBUtil}(\mathcal{I}_1 \mid \mathcal{I}_2) = \mathtt{RBUtil}(\mathcal{I}_1 \cap \mathcal{I}_2) - \mathtt{RBUtil}(\mathcal{I}_2) \ ,$$

and thus we also obtain

$$\mathtt{RBUtil}\big(\mathbb{E}_P[\boldsymbol{\rho}] = \boldsymbol{\mu} \ \big| \ P[\mathcal{E}] = 1\big) = \mathtt{Util}(\boldsymbol{\mu} \mid \mathcal{E}; \boldsymbol{q}) \ .$$

Grünwald and Dawid [16] show that whenever the set $\mathcal{I}$ is closed and convex, the robust Bayes utility $\mathtt{RBUtil}(\mathcal{I})$ coincides with the dual concept of the *maximum (generalized) entropy*, which seeks to find the distribution of the maximum entropy that satisfies a given set of constraints (expressed as $\mathcal{I}$). We do not go into details here, but simply point out that the correspondence between the utility of information and the Bregman divergence (Theorem 1) is just a special case of the duality between the robust Bayes and the maximum entropy.

G SUFFICIENT CONDITIONS AND ROOF EXAMPLES

Here we explore when we can and cannot achieve implicit submarket closing, i.e., ZEROUTIL, EXUTIL, and CONDPRICE simultaneously, in the sudden revelation setting. We begin with an example in which implicit submarket closing is not possible, and then present sufficient conditions, followed by additional examples.

G.1 IMPOSSIBILITY EXAMPLE

Example 9. Consider the square market introduced in Example 2 with the observation function $X(\boldsymbol{\omega}) = \omega_1 + \omega_2 \in \{0, 1, 2\}$. We will see that for this market, the condition of Theorem 2 cannot be satisfied and therefore we cannot achieve EXUTIL. Specifically, we show that there exists an $\boldsymbol{s}$ for which no convex function is consistent with $R^{\hat{\boldsymbol{b}}}$.

First note that the observation function gives rise to conditional price spaces $\mathcal{M}^0 = \{(0,0)\}$, $\mathcal{M}^1 = \mathrm{conv}\{(1,0), (0,1)\} = \{(\lambda, 1-\lambda) : \lambda \in [0,1]\}$, and $\mathcal{M}^2 = \{(1,1)\}$. We examine the value of $R^{\hat{\boldsymbol{b}}}$ at three points,

$$\boldsymbol{\mu}^0 = (0,0) \ , \quad \boldsymbol{\mu}^1 = (\tfrac{1}{2}, \tfrac{1}{2}) \ , \quad \boldsymbol{\mu}^2 = (1,1) \ .$$

By Proposition 2, we have $\hat{b}^x = C(\boldsymbol{s}) - C^x(\boldsymbol{s})$, and so

$$\begin{aligned}
R^{\hat{\boldsymbol{b}}}(\boldsymbol{\mu}^0) &= R(0,0) - [C(\boldsymbol{s}) - C^0(\boldsymbol{s})] = -C(\boldsymbol{s}), \\
R^{\hat{\boldsymbol{b}}}(\boldsymbol{\mu}^2) &= R(1,1) - [C(\boldsymbol{s}) - C^2(\boldsymbol{s})] = -C(\boldsymbol{s}) + s_1 + s_2, \\
R^{\hat{\boldsymbol{b}}}(\boldsymbol{\mu}^1) &= R(\tfrac{1}{2}, \tfrac{1}{2}) - [C(\boldsymbol{s}) - C^1(\boldsymbol{s})] \\
&= -2\ln 2 - C(\boldsymbol{s}) + 2\ln\big(e^{s_1/2} + e^{s_2/2}\big) \\
&= -2\ln 2 - C(\boldsymbol{s}) \\
&\quad + 2\ln\Big[e^{(s_1+s_2)/4}\Big(e^{(s_1-s_2)/4} + e^{(s_2-s_1)/4}\Big)\Big] \\
&= -C(\boldsymbol{s}) + \tfrac{s_1+s_2}{2} + 2\ln\big(\tfrac{z+z^{-1}}{2}\big),
\end{aligned}$$

where $z = e^{(s_1-s_2)/4}$. Note that $\boldsymbol{\mu}^1 = (\boldsymbol{\mu}^0 + \boldsymbol{\mu}^2)/2$, but $R^{\hat{\boldsymbol{b}}}(\boldsymbol{\mu}^1) > (R^{\hat{\boldsymbol{b}}}(\boldsymbol{\mu}^0) + R^{\hat{\boldsymbol{b}}}(\boldsymbol{\mu}^2))/2$ whenever $z + z^{-1} > 2$, i.e., whenever $z > 0$ and $z \neq 1$. From the definition of z this happens whenever $s_1 \neq s_2$, so for any such $\boldsymbol{s}$, no convex function can be consistent with $R^{\hat{\boldsymbol{b}}}$.

G.2 SUFFICIENT CONDITIONS

As we saw in Example 9, there is sometimes tension between satisfying ZEROUTIL and EXUTIL, and in particular, we cannot always achieve both. We now establish *sufficient* conditions under which we can achieve both of these goals (and hence CONDPRICE as well). We will do this in a way that focuses on the geometry of the sets $\mathcal{M}^x$, and consequently our results will apply regardless of the choice of C and the transition state $\boldsymbol{s}$. This not only simplifies the theory, but has practical advantages as well; the market designer need not worry about the transition state, and can choose C independently of concerns about implicit market closing.

In particular, we will show sufficient conditions for when $(\mathrm{conv}\, R^{\hat{\boldsymbol{b}}})$ is consistent with $R^{\hat{\boldsymbol{b}}}$, and then apply Theorem 2. In fact, we show something stronger, by characterizing when $(\mathrm{conv}\, R^{\boldsymbol{b}})$ is consistent with $R^{\boldsymbol{b}}$ for *all* vectors $\boldsymbol{b}$.

Recall that a *face* of a convex set S is a convex subset $F \subseteq S$ such that any line segment in S whose relative interior intersects F, must have both of its endpoints in F. Our sufficient condition requires that the sets $\mathcal{M}^x$ be faces of $\mathcal{M}$. This means that elements of $\mathcal{M}^x$ cannot be obtained as convex combinations including elements from $\mathcal{M}^y$ for $y \neq x$ with non-zero weight.

We define simplices $\Delta_{\mathcal{X}} := \{\boldsymbol{\lambda} \in \mathbb{R}^{\mathcal{X}}_+ : \sum_x \lambda_x \leq 1\}$ and $\Delta_k := \{\boldsymbol{\lambda} \in \mathbb{R}^k_+ : \sum_i \lambda_i \leq 1\}$ where $\mathbb{R}_+$ are non-negative reals. Before proving the sufficient condition, we state the following alternative characterization of the face.

Proposition 6. *Let F and S be convex sets and $F \subseteq S$. Then F is a face of S if and only if for all $\boldsymbol{\mu} \in F$, any decomposition of $\boldsymbol{\mu}$ into a convex combination over S must put zero weight on points outside F; i.e., for all $k \geq 1$, $\boldsymbol{\lambda} \in \Delta_k$ and $\boldsymbol{\mu}^i \in S$ such that $\boldsymbol{\mu} = \sum_{i=1}^k \lambda_i \boldsymbol{\mu}^i$, we must have that $\lambda_i = 0$ for $\boldsymbol{\mu}^i \notin F$.*

Proof. Assume first that F is a face. By convexity of F, a convex combination of any points from F lies in F. Also, any convex combination of points from $S \backslash F$ must lie in $S \backslash F$. This is true for $k = 2$ points by the definition of the face. For $k > 2$ it follows by induction, because, assuming $\lambda_1 > 0$, we can rewrite the convex combination of $\boldsymbol{\mu}_i \in S \backslash F$ as

$$\begin{aligned}
&\lambda_1 \boldsymbol{\mu}_1 + \cdots + \lambda_k \boldsymbol{\mu}_k \\
&\quad = (1 - \lambda_k)\left[\frac{\lambda_1 \boldsymbol{\mu}_1 + \cdots + \lambda_{k-1}\boldsymbol{\mu}_{k-1}}{\lambda_1 + \cdots + \lambda_{k-1}}\right] + \lambda_k \boldsymbol{\mu}_k \ .
\end{aligned}$$

The term in the brackets is in $S \backslash F$ by the inductive hypothesis, so the entire expression is a convex combination of $k = 2$ points from $S \backslash F$, and therefore lies in $S \backslash F$ by the definition of the face. Now assume that $\boldsymbol{\mu} \in F$, and consider any decomposition of $\boldsymbol{\mu}$ into a convex combination over S. By the above reasoning, we can collect the terms with $\boldsymbol{\mu}^i \in F$ and $\boldsymbol{\mu}^i \notin F$ and write $\boldsymbol{\mu} = \lambda_F \boldsymbol{\mu}^F + \lambda_{S\backslash F} \boldsymbol{\mu}^{S\backslash F}$ where λ_F and $\lambda_{S\backslash F}$ are the respective sums of weights of $\boldsymbol{\mu}^i \in F$ and $\boldsymbol{\mu}^i \in S\backslash F$, and $\boldsymbol{\mu}^F \in F$ and $\boldsymbol{\mu}^{S\backslash F} \in S\backslash F$ are their respective convex combinations. From the definition of the face, we obtain $\lambda_{S\backslash F} = 0$.

For the opposite direction, consider any $\boldsymbol{\mu}^1, \boldsymbol{\mu}^2 \in S$ and assume that a point $\boldsymbol{\mu}$ in the relative interior of the connecting line segment lies in F, i.e., $\boldsymbol{\mu} = \lambda_1 \boldsymbol{\mu}^1 + \lambda_2 \boldsymbol{\mu}^2$ with $\lambda_1, \lambda_2 > 0$. The condition of the proposition then implies that the endpoints $\boldsymbol{\mu}^1, \boldsymbol{\mu}^2$ be in F, so F must be a face. ☐

Proposition 7. *For any convex R with $\operatorname{dom} R = \mathcal{M}$, $(\operatorname{conv} R^{\boldsymbol{b}})$ is consistent with $R^{\boldsymbol{b}}$ for all $\boldsymbol{b} \in \mathbb{R}^{\mathcal{X}}$ if and only if the sets $\mathcal{M}^x$ are disjoint faces of $\mathcal{M}$.* [5]

Proof. Suppose that the sets $\mathcal{M}^x$ are disjoint faces of $\mathcal{M}$, and R and $\boldsymbol{b}$ are given. By Proposition B.2.5.1 of Hiriart-Urruty and Lemaréchal [30], we may use an alternate representation of the convex roof,

$$(\operatorname{conv} R^{\boldsymbol{b}})(\boldsymbol{\mu}) = \inf \left\{ \sum_{i=1}^{k} \lambda_i R^{\boldsymbol{b}}(\boldsymbol{\mu}^i) :\ k \geq 1,\ \boldsymbol{\mu}^i \in \mathcal{M}^\star, \boldsymbol{\lambda} \in \Delta_k,\ \sum_{i=1}^{k} \lambda_i \boldsymbol{\mu}^i = \boldsymbol{\mu} \right\}.$$

Intuitively, this expression examines all upper bounds imposed by the convexity constraints from $R^{\boldsymbol{b}}$ and defines $(\operatorname{conv} R^{\boldsymbol{b}})$ as the infimum of these upper bounds. Note that $R^{\boldsymbol{b}}$ is convex on each of the sets $\mathcal{M}^x$ (since it is just a shifted copy of R on $\mathcal{M}^x$). Therefore, we may condense convex combinations within each $\mathcal{M}^x$ (which only lowers the corresponding $R^{\boldsymbol{b}}$ values), yielding

$$(\operatorname{conv} R^{\boldsymbol{b}})(\boldsymbol{\mu}) = \inf \left\{ \sum_{x \in \mathcal{X}} \lambda_x R^{\boldsymbol{b}}(\boldsymbol{\mu}^x) :\ \boldsymbol{\mu}^x \in \mathcal{M}^x, \boldsymbol{\lambda} \in \Delta_{\mathcal{X}},\ \sum_{x \in \mathcal{X}} \lambda_x \boldsymbol{\mu}^x = \boldsymbol{\mu} \right\}. \quad (34)$$

For a given $y \in \mathcal{X}$, the set $\mathcal{M}^y$ is a face disjoint from all $\mathcal{M}^x$ for $x \neq y$. Thus, if $\boldsymbol{\mu} \in \mathcal{M}^y$, we obtain by Prop. 6 that the $\boldsymbol{\lambda}$ in the right hand side of Eq. (34) must have $\lambda_x = 0$ for $x \neq y$ and $\lambda_y = 1$. This immediately yields $(\operatorname{conv} R^{\boldsymbol{b}})(\boldsymbol{\mu}) = R^{\boldsymbol{b}}(\boldsymbol{\mu})$.

For the other direction, first note that if sets $\mathcal{M}^x$ are not disjoint then $R^{\boldsymbol{b}}$ is not well defined for all $\boldsymbol{b}$ and the theorem holds. Assume that sets $\mathcal{M}^x$ are disjoint, but they are not all faces. Therefore, for some $y \in \mathcal{X}$, we have $\boldsymbol{\mu} \in \mathcal{M}^y$ which can be written as a convex combination $\boldsymbol{\mu} = \lambda_1 \boldsymbol{\mu}^1 + \lambda_2 \boldsymbol{\mu}^2$ with $\lambda_1, \lambda_2 > 0$, $\boldsymbol{\mu}^1, \boldsymbol{\mu}^2 \in \mathcal{M}$, but $\boldsymbol{\mu}^1 \notin \mathcal{M}^y$. We will argue that this implies that $\boldsymbol{\mu}$ can be written as a convex combination across $\boldsymbol{\mu}^x \in \mathcal{M}^x$, putting non-zero weight on some $\boldsymbol{\mu}^z$ where $z \neq y$. The reasoning is as follows. Since $\boldsymbol{\mu}^1, \boldsymbol{\mu}^2 \in \mathcal{M}$, they can be written as convex combinations of $\boldsymbol{\rho}(\omega)$ across $\omega \in \Omega$. Collecting $\omega \in \Omega^x$ for $x \in \mathcal{X}$, vectors $\boldsymbol{\mu}^1$ and $\boldsymbol{\mu}^2$ can be in fact written as convex combinations

$$\boldsymbol{\mu}^1 = \sum_{x \in \mathcal{X}} \lambda_{1,x} \boldsymbol{\mu}^{1,x} \quad , \quad \boldsymbol{\mu}^2 = \sum_{x \in \mathcal{X}} \lambda_{2,x} \boldsymbol{\mu}^{2,x}$$

where $\boldsymbol{\mu}^{1,x}, \boldsymbol{\mu}^{2,x} \in \mathcal{M}^x$. Collecting the matching terms, we can thus write $\boldsymbol{\mu}$ as

$$\boldsymbol{\mu} = \sum_{x \in \mathcal{X}} \lambda_x \boldsymbol{\mu}^x$$

where $\lambda_x = \lambda_1 \lambda_{1,x} + \lambda_2 \lambda_{2,x}$ and

$$\boldsymbol{\mu}^x = \frac{\lambda_1 \lambda_{1,x} \boldsymbol{\mu}^{1,x} + \lambda_2 \lambda_{2,x} \boldsymbol{\mu}^{2,x}}{\lambda_1 \lambda_{1,x} + \lambda_2 \lambda_{2,x}} \in \mathcal{M}^x \quad .$$

Since $\boldsymbol{\mu}^1 \notin \mathcal{M}^y$, we must have $\lambda_{1,y} < 1$, and thus also $\lambda_y < 1$ (because $\lambda_1 > 0$). Hence, there must exist some $z \neq y$ such that $\lambda_z > 0$.

To show that $(\operatorname{conv} R^{\boldsymbol{b}})$ cannot be consistent with $R^{\boldsymbol{b}}$ for all $\boldsymbol{b}$, consider $\boldsymbol{b}$ with $b^x = 0$ for $x \neq z$ and b^z equal to some large value. Thus, $\sum_x \lambda_x R^{\boldsymbol{b}}(\boldsymbol{\mu}^x) = \sum_x \lambda_x R(\boldsymbol{\mu}^x) - \lambda_z b^z$. We may make this expression as low as desired by increasing b^z, and in particular, for a sufficiently large b^z, we have $\sum_x \lambda_x R^{\boldsymbol{b}}(\boldsymbol{\mu}^x) < R(\boldsymbol{\mu}) = R^{\boldsymbol{b}}(\boldsymbol{\mu})$, so any function which is consistent with $R^{\boldsymbol{b}}$ will not be convex. ☐

Combining Theorem 2 and Proposition 7, we have the following theorem.

Theorem 10. *If the sets $\mathcal{M}^x$ are disjoint faces of $\mathcal{M}$, then* CondPrice, ExUtil, *and* ZeroUtil *are achieved with* `NewState` *as the identity and* `NewCost` *outputting the conjugate of $\tilde{R} = (\operatorname{conv} R^{\hat{\boldsymbol{b}}})$.*

G.3 BINARY-PAYOFF LCMMS AND THE SIMPLEX

Two key examples studied in this paper are the LMSR on the simplex and LCMMs. In this section, we show that the sufficient condition introduced in the previous section holds for LCMMs with binary payoffs when the payoffs of one submarket are observed, as well as for any observations on a simplex.

[5] If $R^{\boldsymbol{b}}$ is not well defined, we assume that no function can be consistent with $R^{\boldsymbol{b}}$.

We will argue by Theorem 10, showing that the sets $\mathcal{M}^x$ are exposed faces of $\mathcal{M}$. Recall that F is an *exposed face* of a convex set S if F is the set of maximizers of some linear function over S. The exposed face is always a face [27, page 162]

Instead of working with $\mathcal{M}^x$, it in fact suffices to work with Ω^x. Inspired by the definition of an exposed face, we define an "exposed event" as follows.

Definition 9. *An event $\mathcal{E} \subseteq \Omega$ is called* exposed *if it is the set of maximizers of some linear function of $\boldsymbol{\rho}(\omega)$, i.e., if there exists a vector $\boldsymbol{v} \in \mathbb{R}^K$ such that*

$$\mathcal{E} = \operatorname*{argmax}_{\omega \in \Omega} \left[\boldsymbol{v} \cdot \boldsymbol{\rho}(\omega)\right] \quad .$$

It is immediate that if Ω^x is an exposed event, then $\mathcal{M}^x$ is an exposed face disjoint from $\mathcal{M}^y$ for any $y \neq x$. Combining this with Theorem 10 yields the following theorem.

Theorem 11. *If all events Ω^x are exposed, then* CONDPRICE, EXUTIL, *and* ZEROUTIL *are achieved with* `NewState` *as the identity and* `NewCost` *outputting the conjugate of $\tilde{R} = (\mathrm{conv}\, R^{\hat{\boldsymbol{b}}})$.*

We next show how Theorem 11 can be used to argue that submarket closing is possible in binary-payoff LCMMs and on a simplex.

Example 10. *Submarket closing in binary-payoff LCMMs.* We need to argue that the events corresponding to submarket observations in a binary-payoff ($\boldsymbol{\rho}(\omega) \in \{0,1\}^K$ for all $\omega \in \Omega$) LCMMs are exposed. We use the same construction as in the proof of Theorem 7. Let g be a submarket in a binary-payoff LCMM. Let $\boldsymbol{x} \in \mathcal{X}_g$ and $\Omega^{\boldsymbol{x}} := \{\boldsymbol{\rho}_g = \boldsymbol{x}\}$. We need to show that $\Omega^{\boldsymbol{x}}$ is exposed. Consider $\boldsymbol{v} \in \mathbb{R}^K$ with the components

$$v_i = \begin{cases} 1 & \text{if } i \in g \text{ and } x_i = 1, \\ -1 & \text{if } i \in g \text{ and } x_i = 0, \\ 0 & \text{if } i \notin g. \end{cases}$$

Let k be the number of 1s in $\boldsymbol{x}$. Now, as in the proof of Theorem 7, we have $\boldsymbol{v} \cdot \boldsymbol{\rho}(\omega) = k$ for $\omega \in \Omega^{\boldsymbol{x}}$ and $\boldsymbol{v} \cdot \boldsymbol{\rho}(\omega) \leq k - 1$ for $\omega \notin \Omega^{\boldsymbol{x}}$. Thus indeed $\Omega^{\boldsymbol{x}}$ is exposed, and therefore, by Theorem 11, implicit submarket closing is always possible.

Example 11. *Submarket closing on a simplex.* We show that all events on a simplex are exposed and thus any random variable allows implicit submarket closing by Theorem 11. Recall that in a market on a simplex, such as LMSR, we have $\Omega = [K]$ and $\rho_i(\omega) = \mathbf{1}[i = \omega]$. Let $\mathcal{E} \subseteq \Omega$ be an arbitrary event. To see that $\mathcal{E}$ is exposed, consider $\boldsymbol{v} \in \mathbb{R}^K$ with the components $v_i = \mathbf{1}[i \in \mathcal{E}]$. We have

$$\boldsymbol{v} \cdot \boldsymbol{\rho}(\omega) = v_\omega = \mathbf{1}[\omega \in \mathcal{E}] \quad .$$

Thus, $\boldsymbol{v} \cdot \boldsymbol{\rho}(\omega) = 1$ for $\omega \in \mathcal{E}$ and $\boldsymbol{v} \cdot \boldsymbol{\rho}(\omega) = 0$ for $\omega \notin \mathcal{E}$, showing that $\mathcal{E}$ is exposed.

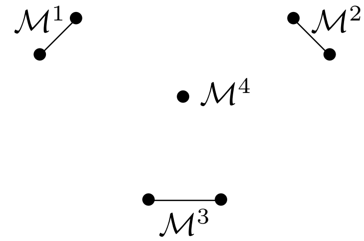

Figure 2: Example showing that the conditions of Theorem 10 are not always necessary.

G.4 WHEN THE SETS $\mathcal{M}^x$ ARE NOT FACES

It is worth noting that the condition in Theorem 10 that requires the sets $\mathcal{M}^x$ to be disjoint faces is merely sufficient and not necessary. In Figure 2 we give a pictorial example in two-dimensional price space in which one of the sets, $\mathcal{M}^4$, is not a face of $\mathcal{M}$, but it is still possible to achieve CONDPRICE, EXUTIL, and ZEROUTIL.

Consider first a market with conditional price spaces $\mathcal{M}^1$, $\mathcal{M}^2$, and $\mathcal{M}^3$ as shown, but *not* $\mathcal{M}^4$. The three sets $\mathcal{M}^1, \mathcal{M}^2$, and $\mathcal{M}^3$ are disjoint faces of $\mathcal{M}$ (the convex hull of these sets), and hence Theorem 10 applies and CONDPRICE, EXUTIL and ZEROUTIL are satisfied by setting the new cost function to the conjugate of $\tilde{R} = (\mathrm{conv}\, R^{(\hat{b}^1, \hat{b}^2, \hat{b}^3)})$. By construction of $\hat{\boldsymbol{b}}$ and $\tilde{R}$, the points $(\hat{\boldsymbol{\mu}}^x, \tilde{R}(\hat{\boldsymbol{\mu}}^x))$ for $x \in \{1, 2, 3\}$ lie on the tangent of R with the slope $\boldsymbol{s}$, and this same hyperplane is also a tangent of $\tilde{R}$ with the slope $\boldsymbol{s}$.

Now consider a market with conditional price spaces $\mathcal{M}^1$, $\mathcal{M}^2$, $\mathcal{M}^3$, and $\mathcal{M}^4$, as in the figure. We will argue that CONDPRICE, EXUTIL, and ZEROUTIL are satisfied for this market using the conjugate of the same function $\tilde{R}$ used above. First observe that the geometry of $\mathcal{M}^1, \mathcal{M}^2, \mathcal{M}^3$, and $\mathcal{M}^4$ implies that regardless of the specific conditional price vectors $\hat{\boldsymbol{\mu}}^x \in \mathcal{M}^x$ for $x \in \{1, 2, 3\}$, we always have that $\hat{\boldsymbol{\mu}}^4$ is in the convex hull of $\hat{\boldsymbol{\mu}}^1$, $\hat{\boldsymbol{\mu}}^2$, and $\hat{\boldsymbol{\mu}}^3$. Now by convexity of $\tilde{R}$, the fact that the tangent to $\tilde{R}$ with slope $\boldsymbol{s}$ contains $(\hat{\boldsymbol{\mu}}^x, \tilde{R}(\hat{\boldsymbol{\mu}}^x))$ for $x \in \{1, 2, 3\}$ implies that this tangent must also contain the point $(\hat{\boldsymbol{\mu}}^4, \tilde{R}(\hat{\boldsymbol{\mu}}^4))$. Thus, setting $b^4 = R(\hat{\boldsymbol{\mu}}^4) - \tilde{R}(\hat{\boldsymbol{\mu}}^4)$, we obtain that $\tilde{R}$ is consistent with $R^{(\hat{b}^1, \hat{b}^2, \hat{b}^3, b^4)}$ (for the same $\hat{b}^1$, $\hat{b}^2$, and $\hat{b}^3$ as above) which by Lemma 2 guarantees CONDPRICE and EXUTIL. Since $(\hat{\boldsymbol{\mu}}^4, \tilde{R}(\hat{\boldsymbol{\mu}}^4))$ is on the tangent, ZEROUTIL holds too.

H BOUNDS ON WORST-CASE LOSS

In this section, we show that the mechanisms studied in this paper maintain an important feature of cost-function-based market makers: a finite bound on the loss of the market maker which is guaranteed to hold no matter what trades are executed or which outcome ω occurs. In particular, we show that the worst-case loss bound of a market maker using the initial cost function (C for sudden revelation market makers, $\mathbf{C}(\cdot;t_0)$ for gradual decrease market makers) is maintained.[6]

For a standard cost-function-based market maker with cost function C, the worst-case market maker loss is simply

$$\texttt{WCLoss}(C;\boldsymbol{s}^{\texttt{ini}}) := \sup_{\omega\in\Omega,\boldsymbol{r}\in\mathbb{R}^K}\left[\boldsymbol{\rho}(\omega)\cdot\boldsymbol{r} - C(\boldsymbol{s}^{\texttt{ini}}+\boldsymbol{r}) + C(\boldsymbol{s}^{\texttt{ini}})\right]$$

where $\boldsymbol{s}^{\texttt{ini}}$ is the initial state of the market. The term inside the supremum is the difference between the amount the market maker must pay traders and the amount collected from traders by the market maker when the cumulative trade vector is $\boldsymbol{r}$ and the outcome is ω. Our assumption that $\operatorname{dom} R = \mathcal{M}$, where R is the conjugate of C, guarantees that $\texttt{WCLoss}(C;\boldsymbol{s}^{\texttt{ini}})$ is always finite [2]. In particular, it is easy to see from Eq. (1) of Theorem 1 that

$$\texttt{WCLoss}(C;\boldsymbol{s}^{\texttt{ini}}) = \max_{\omega\in\Omega} D(\boldsymbol{\rho}(\omega),\boldsymbol{s}^{\texttt{ini}}) \ .$$

We show that the mechanisms introduced in Sections 3 and 4 maintain this bound.

H.1 SUDDEN REVELATION MARKET MAKERS

For sudden revelation market makers (see Protocol 1), the worst-case market maker loss is

$$\begin{aligned}\texttt{WCLoss}(C,\texttt{NewCost},\texttt{NewState};\boldsymbol{s}^{\texttt{ini}}) \\ := \sup_{\omega\in\Omega,\boldsymbol{r}\in\mathbb{R}^K,\tilde{\boldsymbol{r}}\in\mathbb{R}^K}\Big[\boldsymbol{\rho}(\omega)\cdot(\boldsymbol{r}+\tilde{\boldsymbol{r}}) - C(\boldsymbol{s}^{\texttt{ini}}+\boldsymbol{r}) \\ + C(\boldsymbol{s}^{\texttt{ini}}) - \tilde{C}(\tilde{\boldsymbol{s}}+\tilde{\boldsymbol{r}}) + \tilde{C}(\tilde{\boldsymbol{s}})\Big] \quad (35)\end{aligned}$$

where $\tilde{C} = \texttt{NewCost}(\boldsymbol{s}^{\texttt{ini}}+\boldsymbol{r})$ and $\tilde{\boldsymbol{s}} = \texttt{NewState}(\boldsymbol{s}^{\texttt{ini}}+\boldsymbol{r})$. Note that $\tilde{C}$ and $\tilde{\boldsymbol{s}}$ depend on $\boldsymbol{r}$ although we do not write this dependence explicitly. The worst-case loss does not depend on the switch time t.

We now bound this worst case loss for our construction in Sec. 3, with $\tilde{\boldsymbol{s}}$ equal to the state $\boldsymbol{s}$ at the switch time and $\tilde{C}$ defined to be the conjugate of $\tilde{R} = (\operatorname{conv} R^{\hat{\boldsymbol{b}}})$, where $\hat{\boldsymbol{b}}$ depends on $\boldsymbol{s}$. We show that the loss of this market maker is no worse than that of a market maker using the initial cost function C.

Theorem 12. *If* $\texttt{NewState}(\boldsymbol{s}) = \boldsymbol{s}$ *and* $\texttt{NewCost}(\boldsymbol{s})$ *is defined as in Theorem 2, then for any bounded-loss, no-arbitrage cost function* C *and any initial state* $\boldsymbol{s}^{\texttt{ini}}$*,*

$$\texttt{WCLoss}(C,\texttt{NewCost},\texttt{NewState};\boldsymbol{s}^{\texttt{ini}}) \le \texttt{WCLoss}(C;\boldsymbol{s}^{\texttt{ini}}) .$$

Proof. Let $\tilde{\boldsymbol{s}}^{\texttt{fin}}$ be the final state of the market and $\boldsymbol{s}$ be the market state at the switch time t, as in Protocol 1. Then from Proposition 2, $\tilde{C}(\tilde{\boldsymbol{s}}) = \tilde{C}(\boldsymbol{s}) = C(\boldsymbol{s})$ and

$$\begin{aligned}&\texttt{WCLoss}(C,\texttt{NewCost},\texttt{NewState};\boldsymbol{s}^{\texttt{ini}}) \\ &= \max_{\omega\in\Omega}\sup_{\tilde{\boldsymbol{s}}^{\texttt{fin}}\in\mathbb{R}^K}\Big[\boldsymbol{\rho}(\omega)\cdot(\tilde{\boldsymbol{s}}^{\texttt{fin}}-\boldsymbol{s}^{\texttt{ini}}) \\ &\qquad + C(\boldsymbol{s}^{\texttt{ini}}) - \tilde{C}(\tilde{\boldsymbol{s}}^{\texttt{fin}})\Big] \ .\end{aligned}$$

By conjugacy we have

$$\sup_{\tilde{\boldsymbol{s}}^{\texttt{fin}}\in\mathbb{R}^K}\Big[\boldsymbol{\rho}(\omega)\cdot\tilde{\boldsymbol{s}}^{\texttt{fin}} - \tilde{C}(\tilde{\boldsymbol{s}}^{\texttt{fin}})\Big] = \tilde{R}\big(\boldsymbol{\rho}(\omega)\big) \ .$$

By the definition of $\hat{\boldsymbol{b}}$,

$$\tilde{R}\big(\boldsymbol{\rho}(\omega)\big) = R\big(\boldsymbol{\rho}(\omega)\big) - D(\hat{\boldsymbol{\mu}}^x\|\boldsymbol{s}) \le R\big(\boldsymbol{\rho}(\omega)\big)$$

for some $\hat{\boldsymbol{\mu}}^x \in \boldsymbol{p}(\Omega^x;\boldsymbol{s})$ where $x\in\mathcal{X}$ is such that $\omega\in\Omega^x$. Putting this together, we obtain the bound

$$\begin{aligned}&\texttt{WCLoss}(C,\texttt{NewCost},\texttt{NewState};\boldsymbol{s}^{\texttt{ini}}) \\ &\le \max_{\omega\in\Omega}\left[R\big(\boldsymbol{\rho}(\omega)\big) + C(\boldsymbol{s}^{\texttt{ini}}) - \boldsymbol{\rho}(\omega)\cdot\boldsymbol{s}^{\texttt{ini}}\right] \\ &= \texttt{WCLoss}(C;\boldsymbol{s}^{\texttt{ini}}) \ . \qquad \square\end{aligned}$$

H.2 GRADUAL DECREASE LCMMS

For gradual decrease market makers (see Protocol 2), the worst-case market maker loss can be written as

$$\begin{aligned}&\texttt{WCLoss}(\mathbf{C},\texttt{NewState};\boldsymbol{s}^0,t^0) \\ &:= \sup_{\substack{\omega\in\Omega,N\ge 0,\{\boldsymbol{r}^i\}_{i=1}^N,\{t^i\}_{i=1}^N \\ \text{with } t^0\le t^1\le\cdots\le t^N}}\Bigg[\sum_{i=1}^N\Big[\boldsymbol{\rho}(\omega)\cdot\boldsymbol{r}^i \\ &\qquad - \mathbf{C}(\tilde{\boldsymbol{s}}^{i-1}+\boldsymbol{r}^i;t^i) + \mathbf{C}(\tilde{\boldsymbol{s}}^{i-1};t^i)\Big]\Bigg] \quad (36)\end{aligned}$$

where $\tilde{\boldsymbol{s}}^{i-1} = \texttt{NewState}(\boldsymbol{s}^{i-1};t^{i-1},t^i)$.

We next show that the worst-case loss of the gradual decrease LCMM developed in Sec. 4 is no worse than that of a market maker using the initial cost function $\mathbf{C}(\cdot;t^0)$.

Theorem 13. *For the gradual decrease LCMM with corresponding function* $\texttt{NewState}$ *and cost* $\mathbf{C}$ *and any differentiable non-increasing information-utility schedules* β_g*, for any initial state* $\boldsymbol{s}_0$ *and time* t_0*,*

$$\texttt{WCLoss}(\mathbf{C},\texttt{NewState};\boldsymbol{s}^0,t^0) \le \texttt{WCLoss}(C^0;\boldsymbol{s}^0)$$

where $C^0 := \mathbf{C}(\cdot;t^0)$*.*

Proof. In the context of Protocol 2, let C^i denote $\mathbf{C}(\cdot;t^i)$, and R^i and D^i denote the corresponding conjugate and divergence. First, note that by Theorem 3, for any i and any

[6] We actually show something slightly stronger: for every outcome ω, the worst case loss of the market maker conditioned on the true outcome being ω is maintained.

$\boldsymbol{\mu} \in \mathcal{M}$, for suitable $\boldsymbol{\delta}^\star$ and $\boldsymbol{\eta}^\star$,

$$\begin{aligned} & D^{i+1}(\boldsymbol{\mu} \| \tilde{\boldsymbol{s}}^i) \\ & \quad = \sum_{g \in \mathcal{G}} \frac{\beta_g(t^{i+1})}{\beta_g(t^i)} D_g^i(\boldsymbol{\mu}_g \| \boldsymbol{s}_g^i + \boldsymbol{\delta}_g^\star) + (\mathbf{A}^\top \boldsymbol{\mu} - \boldsymbol{b}) \cdot \boldsymbol{\eta}^\star \\ & \quad \leq \sum_{g \in \mathcal{G}} D_g^i(\boldsymbol{\mu}_g \| \boldsymbol{s}_g^i + \boldsymbol{\delta}_g^\star) + (\mathbf{A}^\top \boldsymbol{\mu} - \boldsymbol{b}) \cdot \boldsymbol{\eta}^\star \\ & \quad = D^i(\boldsymbol{\mu} \| \boldsymbol{s}^i) \ . \end{aligned} \tag{37}$$

The last equality follows from Theorem 6c.

We can bound the expression inside the supremum in Eq. (36) as

$$\begin{aligned} & \sum_{i=1}^N \Big[\boldsymbol{\rho}(\omega) \cdot \boldsymbol{r}^i - C^i(\tilde{\boldsymbol{s}}^{i-1} + \boldsymbol{r}^i) + C^i(\tilde{\boldsymbol{s}}^{i-1}) \Big] \\ & \quad = \sum_{i=1}^N \Big[R^i\big(\boldsymbol{\rho}(\omega)\big) + C^i(\tilde{\boldsymbol{s}}^{i-1}) - \boldsymbol{\rho}(\omega) \cdot \tilde{\boldsymbol{s}}^{i-1} \\ & \qquad\qquad - R^i\big(\boldsymbol{\rho}(\omega)\big) - C^i(\tilde{\boldsymbol{s}}^{i-1} + \boldsymbol{r}^i) + \boldsymbol{\rho}(\omega) \cdot (\tilde{\boldsymbol{s}}^{i-1} + \boldsymbol{r}^i) \Big] \\ & \quad = \sum_{i=1}^N \Big[D^i(\boldsymbol{\rho}(\omega) \| \tilde{\boldsymbol{s}}^{i-1}) - D^i(\boldsymbol{\rho}(\omega) \| \tilde{\boldsymbol{s}}^{i-1} + \boldsymbol{r}^i) \Big] \\ & \quad = \sum_{i=1}^N \Big[D^i(\boldsymbol{\rho}(\omega) \| \tilde{\boldsymbol{s}}^{i-1}) - D^i(\boldsymbol{\rho}(\omega) \| \boldsymbol{s}^i) \Big] \\ & \quad = D^1(\boldsymbol{\rho}(\omega) \| \tilde{\boldsymbol{s}}^0) + \sum_{i=1}^{N-1} \Big[D^{i+1}(\boldsymbol{\rho}(\omega) \| \tilde{\boldsymbol{s}}^i) - D^i(\boldsymbol{\rho}(\omega) \| \boldsymbol{s}^i) \Big] \\ & \qquad\qquad - D^N(\boldsymbol{\rho}(\omega) \| \boldsymbol{s}^N) \\ & \quad \leq D^0(\boldsymbol{\rho}(\omega) \| \boldsymbol{s}^0) \end{aligned}$$

where the last inequality follows by applications of Eq. (37) to the first two terms and the positivity of $D^N(\cdot \| \cdot)$. Taking the supremum, we obtain

$$\begin{aligned} & \texttt{WCLoss}(\mathbf{C}, \texttt{NewState}; \boldsymbol{s}^0, t^0) \\ & \quad \leq \max_{\omega \in \Omega} D^0(\boldsymbol{\rho}(\omega) \| \boldsymbol{s}^0) = \texttt{WCLoss}(C^0; \boldsymbol{s}^0) \ . \qquad \square \end{aligned}$$