Computer Science and Artificial Intelligence Laboratory
Technical Report

MIT-CSAIL-TR-2010-059 December 20, 2010

Conservative-Bayesian Mechanism Design

Pablo Azar, Jing Chen, and Silvio Micali

massachusetts institute of technology, cambridge, ma 02139 usa — www.csail.mit.edu

Conservative-Bayesian Mechanism Design*

Pablo Azar
CSAIL, MIT
Cambridge, MA 02139, USA
azar@csail.mit.edu

Jing Chen
CSAIL, MIT
Cambridge, MA 02139, USA
jingchen@csail.mit.edu

Silvio Micali
CSAIL, MIT
Cambridge, MA 02139, USA
silvio@csail.mit.edu

December 20, 2010

Abstract

Classical Bayesian mechanism design is *"centralized"*, that is, the designer is assumed to know the distribution $\mathcal{D}$ from which the players' type profile has been drawn. We instead investigate a very *"decentralized"* Bayesian model, where the designer has no knowledge at all, and each player only has some probabilistic information about $\mathcal{D}$.

For this decentralized model and many contexts of interest, where the goal is to maximize revenue, we show that, *for arbitrary type distributions $\mathcal{D}$ (in particular, correlated ones),* it is possible to design mechanisms matching to a significant extent the performance of the optimal centralized mechanisms.

Our results are "existential" for a broad class of contexts (including combinatorial auctions) and "constructive" for auctions of a single good.

*This work has been partially funded by the Office of Naval Research under award number N00014091059. (Any opinions, findings, and conclusions or recommendations expressed in this email are those of the author (s) and do not necessarily reflect the views of the Office of Naval Research.)

1 Introduction

In a game, the players' utilities are determined by their (utility) types, and each player knows his own type. In most real games, however, the players do not exactly know the types of their opponents. It is thus both traditional and natural to model such uncertainty by means of probabilistic distributions.[1]

At the core of Bayesian mechanism design is the assumption that the true type profile of the players, θ, has been drawn from a distribution $\mathcal{D}$ over Θ, the set of all possible type profiles. But then different assumptions are made about who knows how much about $\mathcal{D}$. Most mechanisms are defined under the following

Centralized-Bayesian Assumption: *The Designer knows $\mathcal{D}$.*

Broadly speaking, we are interested in investigating how well the performance of these mechanisms can be approximated under *decentralized* Bayesian assumptions, that is, when all knowledge about $\mathcal{D}$ is assumed to lie with the players themselves. Such assumptions are more realistic whenever the players "know each other more than the designer knows them" (e.g., when the designer is a judge allocating an inheritance among family members). Some examples of decentralized-Bayesian assumptions are

Decentralized Assumption 1: *$\mathcal{D}$ is common knowledge among the players;* and

Decentralized Assumption 2: *Each player knows $\mathcal{D}$.*

Weaker and more sophisticated decentralized assumptions envisage that each player, rather than knowing $\mathcal{D}$, has only some partial, probabilistic knowledge about $\mathcal{D}$. In particular, this is the case for the following

Decentralized Assumption 3: *Each player i's knowledge coincides with the conditional distribution $\mathcal{D}|\theta_i$.*

That is, i's knowledge coincides with $\mathcal{D}$ conditioned on the event $\{t \in \Theta : t_i = \theta_i\}$. Note that Decentralized Assumption 3 is indeed weaker than Decentralized Assumption 2, because i can deduce $\mathcal{D}|\theta_i$ from his knowledge of $\mathcal{D}$ and θ_i, but not vice-versa, as in principle i may know $\mathcal{D}|\theta_i$ without knowing $\mathcal{D}$ itself. Also note that Decentralized Assumption 3 is the one underlying most Bayesian mechanisms using Bayesian-Nash equilibria as their solution concept. We are actually interested in investigating the power of a weaker decentralized-Bayesian assumption.

1.1 Our Conservative-Bayesian Model

Notice that a Bayesian mechanism designed under Decentralized Assumption 3 works *if and only if* each player's distributional knowledge is so coarse as to coincide precisely with $\mathcal{D}$ conditioned on "his own true type and nothing more." Should a player be able to further refine the conditional distribution $\mathcal{D}|\theta_i$ (e.g., by acquiring additional information about his opponents), he might have all the reasons in the world to deviate from his strategy in a typical Bayesian-Nash equilibrium. We thus wish to design decentralized-Bayesian mechanisms that achieve their desiderata no matter how more refined the distributional knowledge of the players may be. Let us explain.

Call a partition $\mathcal{P}$ of Θ *i-consistent* if for any set S in $\mathcal{P}$ and any type profiles t and t' in S, $t_i = t'_i$.

Conservative-Bayesian Assumption: There exists a profile $\mathscr{P}$ of partitions of Θ such that: (1) $\mathscr{P}_i$ is i-consistent for each i, and (2) when θ is randomly selected according to $\mathcal{D}$, each player i's knowledge is $\mathcal{D}'_i \triangleq \mathcal{D}|S_i$, where S_i is the unique set in $\mathscr{P}_i$ containing θ.

That is, each $\mathcal{D}'_i$ is a *"separate and arbitrary refinement of $\mathcal{D}|\theta_i$".* For concreteness we only assume that each $\mathcal{D}'_i$ is described by *finitely many* bits. Note that our conservative-Bayesian assumption

(a) *makes no restriction on $\mathcal{D}$*: indeed, $\mathcal{D}$ may be any member of $\Delta(\Theta)$, the set of all distributions over Θ (in particular, the players' types can be arbitrarily correlated);

[1] In the most general Bayesian setting, each player has his own, arbitrary *belief*, that is, he considers the types of his opponents drawn from a probability distribution $\mathcal{D}_i$. Such beliefs may be "unrelated to the truth," and beliefs of different players may be inconsistent with each other. It is hard, however, to design mechanisms that can guarantee desirable outcomes in such a setting. In particular, it would be helpful for the players' beliefs to be consistent with the truth, and thus with each other.

(b) *is totally decentralized:* indeed the designer may know (arbitrarily little or even) nothing about $\mathcal{D}$; and

(c) *is weaker than other decentralized assumptions:* in particular, it is weaker than Assumptions 1, 2, and 3.

1.2 Assignment Contexts

Recall that a game consists of a *context*, describing the outcomes and the players' types, knowledge, and utilities, and a *mechanism*, describing the strategies available to the players and how these strategies lead to outcomes. We focus on a class of contexts general enough to include as special cases all types of auctions, from single-good to combinatorial. We refer to this class as "assignment contexts." In essence, an assignment is an allocation whose different components may overlap, and in our contexts —as in typical auctions— the set of possible assignments is "downward-closed." More formally,

An *assignment context* consists of

- A finite set of *players* $N = \{1, \ldots, n\}$.
- A finite set X of *(assignable) items.*
- A set of *(possible) assignments*: $\mathbb{A} \subset (2^X)^n$ such that $(A_1, \ldots, A_{i-1}, \emptyset, A_{i+1}, \ldots, A_n) \in \mathbb{A}\ \forall A \in \mathbb{A}, i \in N$.
- A profile Θ of *possible types*, where each Θ_i is the set of all functions t from 2^X to $\mathbb{Z}^+$ such that $t(\emptyset) = 0$.
- The set of *outcomes*, $\Omega = \mathbb{A} \times \mathbb{R}^n$.

The utility of a player i in an outcome (A, P) is $\theta_i(A_i) - P_i$.

For finiteness, we assume that each type t actually maps 2^X to $[B] = \{0, \ldots, B-1\}$, for some integer B.

1.3 Our Results

Informal Notation

- We call a mechanism "classical DST" if, for each player i, i's strategy space is Θ_i and it is dominant-strategy for him to announce his own true type.
- We call a mechanism "two-step DST" if, for each player i, i's strategy space is $\Theta_i \times \Delta(\Theta)$ (i.e., his utility type and his knowledge) and i maximizes his own utility by (1) announcing his own true utility type regardless of the strategies of the others; and (2) announcing his true knowledge, given that all the other players announce their own true utility types. (Note that "two-step DST does not imply normal-form.")
- Let M be a decentralized-Bayesian mechanism. Then, by $\text{REV}(M, C, \mathcal{D})$ and $\text{OPT}(C, \mathcal{D})$ we denote the expected revenue —in an execution where the type-distribution is $\mathcal{D}$, the actual context C, and each player is truthful— respectively generated by M and the optimal classical DST centralized-Bayesian mechanism for $\mathcal{D}$. (Each expectation is taken over all possible random choices, that is, over $\mathcal{D}$ and the coin tosses of the corresponding mechanism, if probabilistic.)

1.3.1 An Existential Result for General Assignments

We show that conservative-Bayesian Mechanisms are in principle as powerful as centralized-Bayesian ones.

Theorem 1. *For any $\epsilon > 0$, there exists a two-step DST conservative-Bayesian mechanism $\mathbb{M}$ such that, for all n-player assignment contexts C and type-distributions $\mathcal{D}$,*

$$\text{REV}(\mathbb{M}, C, \mathcal{D}) > (1 - \frac{1}{n})\text{OPT}(C, \mathcal{D}) - \epsilon.$$

While quite general, we consider this result "existential" because our proof is non-constructive. That is, we always guarantee the *existence* of the required $\mathbb{M}$, but cannot always explicitly *construct* it.

1.3.2 A Constructive Result for Single-Good Auctions with Arbitrary Type-Distributions

For single-good auction contexts, we *constructively* provide a lower bound to the power of conservative-Bayesian mechanism design, *for all possible type-distributions*, in terms of the following benchmark.

The Revenue Benchmark $\mathscr{S}$ In a single-good auction with (arbitrary) type-distribution $\mathcal{D}$ and realized type-profile θ, letting the *star player* —denoted by $\star$— be the lexicographically first player having the highest valuation for the good, the *star benchmark*, $\mathscr{S}$, is informally defined to consist of the maximum expected revenue obtainable by

1. *Choosing a price $p_\star$ given inputs $\mathcal{D}$, $\star$, and $\theta_{-\star}$, and*
2. *Collecting payment $p_\star$ whenever $p_\star \leq \theta_\star$.*

(For a formalization of $\mathscr{S}$ see Section 3.) For this benchmark we prove the following

Theorem 2. *For any $\epsilon, \delta > 0$ there exists an* `explicit`*, probabilistic, two-step DST, conservative-Bayesian mechanism $\mathscr{M}$ such that, for all single-good auction contexts C and type-distributions $\mathcal{D}$,*

$$\text{REV}(\mathscr{M}, C, \mathcal{D}) \geq (1-\delta)\mathscr{S} - \epsilon.$$

1.3.3 Constructive Comparisons with Optimal Mechanisms

Theorems 1 and 2 have non-trivial implications for the relative performance of constructive conservative-Bayesian mechanisms and optimal ones in each of three possible scenarios for auctions of a single good.

1. Players with Arbitrarily Correlated Types

When $\mathcal{D}$ is an arbitrary *joint* distribution, no optimal centralized-Bayesian mechanisms are known, and Myerson's mechanism [11] does not guarantee any significant revenue. Yet, Theorem 2 implies the following

Corollary 1. *For any $\epsilon, \delta > 0$ there exists an* `explicit`*, two-step DST, conservative-Bayesian mechanism M such that for all single-good auction contexts C and type-distributions $\mathcal{D}$,*

$$\text{REV}(M, C, \mathcal{D}) \geq \frac{1-\delta}{2}\text{OPT}(C, \mathcal{D}) - \epsilon$$

Indeed, Ronen [13] proves that, for the single-good case, Vickrey auctions with monopoly reserve prices give at least $\frac{1}{2}$ of the optimal revenue, even when valuations are allowed to be correlated. In our parlance: $\mathscr{S} \geq \text{OPT}(C, \mathcal{D}))/2$ for all C and $\mathcal{D}$.

2. Players with Independently Distributed Types

In auctions of a single good where $\mathcal{D}$ is a *product distribution,* that is, when $\mathcal{D} = D_1 \times \cdots \times D_n$, the optimal, classical DST, centralized-Bayesian mechanism is well known: namely, it is Myerson's mechanism [11].

Theorem 1 automatically guarantees that the performance of Myerson's mechanism can be essentially matched by a conservative-Bayesian mechanism $\mathbb{M}$. As stated, Theorem 1 only guarantees the *existence* of $\mathbb{M}$. But our proof of Theorem 1 actually guarantees that $\mathbb{M}$ can be *explicitly* constructed for *all* type-distributions $\mathcal{D}$ for which the optimal, classical DST, centralized-Bayesian mechanism has been explicitly constructed. In the case at hand, therefore, our proof of Theorem 1 immediately yields the following

Corollary 2. *For any $\epsilon > 0$, there exists an* `explicit`*, two-step DST, conservative-Bayesian mechanism M' such that, for all n-player single-good auction contexts C and all product distributions $\mathcal{D} = D_1 \times \cdots \times D_n$,*

$$\text{REV}(M', C, \mathcal{D}) > (1 - \frac{1}{n})\text{OPT}(C, \mathcal{D}) - \epsilon.^2$$

[2] We note that our result more generally applies to all "single-parameter downward-closed environments."

Our small degradation in performance may perhaps be excused. After all, Myerson's mechanism is *centralized*, and thus has access to the product distribution $\mathcal{D}$ for free, while our M' is *very decentralized*, and thus sacrifices some revenue in order to "extract" $\mathcal{D}$ from the "collective" knowledge of the players.

To be sure, other explicit and decentralized-Bayesian mechanisms have already been proposed for single-good auctions, but only in the more specialized scenario discussed below.

3. Players with Identically and Independently Distributed Types

In single-good auctions where $\mathcal{D}$ is "iid", that is, when $\mathcal{D} = D \times \cdots \times D$, explicit and decentralized-Bayesian mechanisms have been proposed by Segal [14] and Baliga and Vohra [1]. Informally speaking, their mechanisms estimate $\mathcal{D}$ using the valuations reported by the players and then run Myerson's mechanism. As the number of players goes to infinity, the estimated distribution becomes closer to $\mathcal{D}$, and the auction's revenue approximates the optimal one. Baliga and Vohra also make a similar analysis for double auctions, where the auctioneer is a broker that matches buyers and sellers.

As already said, our explicit mechanism M' of Corollary 2 is two-step DST and does not need the players' types to be iid. Yet, when restricting our attention to the latter case, a main difference separates our mechanism from theirs. Namely, their mechanisms approximate well the optimal revenue *asymptotically* but not, for *all values* of n. By contrast, the revenue of M' is *always* arbitrarily close to a fraction $\frac{n-1}{n}$ of the optimal revenue.

1.4 Computation and Techniques

Although computational efficiency is not the primary goal of this paper, we note that our mechanisms M and M' of Corollaries 1 and 2 are always computationally tractable when the players' valuations have small range, an important setting in which our mechanisms still retain their advantages over all previous ones. In addition, if we were just content to guarantee our players positive utility in expectation rather than "in all possible cases", then

- M is always computationally tractable, and
- If the optimal mechanism is computationally tractable, or has a computationally tractable approximation satisfying some mild technical conditions, then M' too is always computationally tractable.

Our techniques are conceptually simple. In essence, we integrate Vickrey auctions and scoring rules (a technique from statistics, mostly applied in prediction markets). Although overlooked, this is a powerful integration, and we believe and hope that it will enable all of us to reach many other desirable goals.

2 Other Related Work

Attribute-Based Mechanisms For downward-closed single-parameter contexts, Dhangwatnotai, Roughgarden, and Yan [5] show how to obtain approximately optimal revenue when (1) the players are assumed to be described by some attribute a, (2) all players with this attribute have a one-dimensional valuation drawn from a distribution $\mathscr{D}_a$, and (3) for every attribute a there exist at least two players described by a.

Prior-Free Mechanisms in Digital-Good Auctions Goldberg, Hartline, Karlin, Saks and Wright [8] consider non-Bayesian auctions of digital goods. (A good g is called "digital" if an unlimited number of copies of g can be generated at no additional cost, and the value that each player may have for any copy of g is the same.) For such auctions, they put forward a DST mechanism whose expected revenue, for any possible type profile θ, is guaranteed to be at least a fraction $\frac{1}{4}$ of the following benchmark: $\mathcal{F}^2(\theta) = \max_{i \geq 2} i \cdot \theta_{(i)}$ where $\theta_{(i)}$ is the i^{th} highest valuation in θ. (Since their context is not Bayesian, the expectation is taken solely over the mechanism's coin tosses.) Using this framework, Goldberg and Hartline [7] propose a quite different mechanism, again for auctions of digital goods, achieving a fraction $\frac{1}{3.39}$ of the same benchmark.

Simple Mechanisms Neeman [12] shows that English Auctions are approximately optimal, performing an analysis similar to Ronen's work on Vickrey auctions. Hartline and Roughgarden [10] extend Ronen's results to downward-closed and matroid environments, under the assumptions that the distribution of valuations are independent and satisfy some regularity properties.

Posted-Price mechanisms Chawla, Hartline, Malec and Sivan [3] show a sequential-posted price mechanism that achieves a constant-factor approximation of the optimal revenue even in multiple-parameter settings. In particular, they consider the case where agents desire only one good, but may have different valuations for different goods.

3 Preliminaries

3.1 Our Star Benchmark

In a single-good context with an arbitrary type-distribution $\mathcal{D}$, the star benchmark $\mathscr{S}$ consists of

$$\sum_{i=1}^{n} \sum_{t \in [B]^{n-1}} \Pr_{\theta \leftarrow \mathcal{D}}(\star = i, \theta_{-\star} = t) \max_p \left(p \cdot \Pr_{\theta \leftarrow \mathcal{D}}(\theta_\star \geq p | \star = i, \theta_{-\star} = t_{-i}) \right).$$

3.2 Our Two Building Blocks

We describe our mechanisms in a modular way using the following two building blocks.

Knowledge Aggregator. We define the *knowledge aggregator,* `AGG`, to be the function mapping the identity of a player i and a distribution subprofile $\mathcal{DK}_{-i}$ to another distribution as follows.

`AGG`$(i, \mathcal{DK}_{-i})$

0. For each $j \neq i$, set S_j to be the support of $\mathcal{DK}_j$.
1. Set $S = \cap_{j \neq i} S_j$ and $j' = \min\{j : j \neq i\}$.
2. If $S = \emptyset$, then output $\mathcal{DK}_{j'}$.
3. Else, set $\mathcal{DK}'$ to be $\mathcal{DK}_{j'}|S$, and output $\mathcal{DK}'$.

In essence, `AGG` interprets each $\mathcal{DK}_j$ as the distributional knowledge of player j in a conservative-Bayesian model with type-distribution $\mathcal{D}$, and aggregates the individual knowledge of the players in $-i$ so as to reconstruct a refined distribution for i's type. Notice that in `AGG`, if the knowledge of players in $-i$ disagrees with each other in an apparent way (that is, when $S = \emptyset$), then the function's output can actually be arbitrary. Also notice that the choice of the player j' can be arbitrary, because when the input $\mathcal{DK}$ is the "true knowledge" of the players, any choice of j' will lead to the same output $\mathcal{DK}'$. Here by "true knowledge" we mean that, each $\mathcal{DK}_j$ is obtained from the true distribution $\mathcal{D}$ by first conditioning on some event that is consistent with the true valuation profile, and then conditioning on the true valuation subprofile of players in $-i$.

Brier's Scoring Rule [2]. Let Ω be a state space and let $\Delta(\Omega)$ be the set of probability distributions over Ω. A scoring rule S is a function, $S : \Omega \times \Delta(\Omega) \to \mathbb{R}$. A scoring rule S can be used to reward individuals reporting their knowledge (or beliefs) about the world: in particular, giving an individual reporting a probability distribution $\mathcal{D} \in \Delta(\Omega)$ a reward equal to $S(\omega, \mathcal{D})$ whenever the realized state is $\omega \in \Omega$. A scoring rule S is *proper* if an individual maximizes his expected reward by announcing his true knowledge about the world. That is, for any two different probability distributions $\mathcal{D}, \mathcal{P} \in \Delta(\Omega)$,

$$\mathbb{E}_{\omega \leftarrow \mathcal{D}}[S(\omega, \mathcal{D})] \geq \mathbb{E}_{\omega \leftarrow \mathcal{D}}[S(\omega, \mathcal{P})].$$

A recent paper by Gneiting and Raftery is a good survey of proper scoring rules [6]. We call scoring rule S *strictly proper* if the above inequality is strict. In our mechanisms we use Brier's scoring rule [2] for discrete domains, which is strictly proper. This scoring rule, denoted by BSR, is defined as follows. For each $s \in \Omega$, letting $\mathcal{D}(s)$ be the probability assigned to s by $\mathcal{D}$, and letting $\delta_{\omega,s}$ be the indicator function, that is $\delta_{\omega,s} = 1$ if $s = \omega$ and 0 otherwise, then

$$\boxed{\mathtt{BSR}(\omega, \mathcal{D}) = -\left(\sum_{s\in\Omega}(\delta_{\omega,s} - \mathcal{D}(s))^2\right) = 2\mathcal{D}(\omega) - \|\mathcal{D}\|_2^2 - 1.}$$

Note that Brier's scoring rule is always *bounded:* indeed, $\mathtt{BSR}(\omega, \mathcal{D}) \in [-2, 0]$ for all ω and $\mathcal{D}$. (In contrast, Good's [9] more popular *logarithmic scoring rule* $\mathtt{LSR}(\omega, \mathcal{D}) = \log(\mathcal{D}(\omega))$ is unbounded.)

4 Proof of Theorem 1

Fixing $\epsilon > 0$ we now prove the existence of the required conservative-Bayesian mechanism $\mathbb{M}$. Our $\mathbb{M}$ first obtains from a player a distribution over the types of the other players, and then runs the optimal centralized-Bayesian, DST mechanism for this $(n-1)$-player distribution. Since the latter mechanism is in general unknown, our proof is "non-constructive." But it is a valid "existential" one, because in the worst case, such mechanism can always be found via an exhaustive search in a space that is finite by definition. (In essence our construction of $\mathbb{M}$ is a *reduction*, explicit whenever the optimal mechanism is explicitly known.)

In our description below, numbered steps are taken by the players, and steps marked by letters are steps taken by the mechanism/auctioneer.

Mechanism $\mathbb{M}$

a. *Choose a player i uniformly at random from $\{1, ..., n\}$.*
Comment. Player i will receive the empty allocation and get a price of zero, but he will be rewarded according to his knowledge. Choosing i deterministically does not affect incentives, but might reduce revenue. We do not know how to make $\mathbb{M}$ two-step DST without "removing one player".

1. *Player i announces a distribution $\mathcal{F}$ over Θ_{-i}.*
Comment. Allegedly, $\mathcal{F}$ represents his true knowledge $\mathcal{D}'_i$ restricted to the domain Θ_{-i}.

2. *Each player $j \neq i$ announces a valuation function $v_j \in \Theta_j$.*
Comment. Allegedly, v_j is j's true valuation.

b. *Letting $\mathcal{M}$ be the optimal mechanism for the $(n-1)$-player distribution $\mathcal{F}$, run $\mathcal{M}(v_{-i}, \mathcal{F})$ so as to obtain an allocation $A_{-i} = (A_1, A_2, ..., A_{i-1}, A_{i+1}, ..., A_n)$, and a price vector $P_{-i} = (P_1, P_2, ..., P_{i-1}, P_{i+1}, ...P_n)$.*
Comment. The allocations and prices of the players in $-i$ are determined by $\mathcal{M}$.

c. *Set $A_i = \emptyset$, reward i by setting $P_i = -\frac{\epsilon}{2}(2 + \mathtt{BSR}(v_{-i}, \mathcal{F}))$, and output outcome (A, P).*
Comment. Although player i gets no allocation, he gets a reward according to Brier's scoring rule. The negative price indicates that the mechanism transfers money to player i. Since we insist that our players do not incur negative utilities *ex post*, the boundedness of Brier's scoring rule is crucial here.

Notation. Since a direct (possibly randomized) centralized-Bayesian assignment mechanism M may be conceptualized as receiving (the description of) a type-distribution $\mathcal{D}$ as a separate input, we use a slightly different notation to denote M's revenue than the one used for decentralized-Bayesian mechanisms. Namely, if M always flips ℓ coins, then letting v be a type profile and $r \in \{0,1\}^\ell$, $M(v, \mathcal{D}, r)$ denotes the unique outcome (A, P) computed by M on inputs v and $\mathcal{D}$ when using coin tosses r; $\text{REV}(M(v, \mathcal{D}, r)) = \sum_{i=1}^n P_i$; $\text{REV}(M(v, \mathcal{D})) = \mathbb{E}_{r\leftarrow\{0,1\}^\ell}\text{REV}(M(v, \mathcal{D}, r))$; and $\text{REV}(M(\mathcal{D})) = \mathbb{E}_{v\leftarrow\mathcal{D}}\mathbb{E}_{r\leftarrow\{0,1\}^\ell}\text{REV}(M(v, \mathcal{D}, r))$.

Lemma 1. *The mechanism $\mathbb{M}$ is two-step DST.*

Proof. Let i be the player chosen in step a of the mechanism. First we prove that each player $j \neq i$ maximizes his utility by announcing his true valuation in step 2, regardless of the other players' strategies. Indeed, player j gets allocation A_j and price P_j, which are determined by running the optimal classical-DST mechanism $\mathcal{M}(v_{-i}, \mathcal{F})$. Since $\mathcal{M}$ is classical DST, player j maximizes his utility $\theta_j(A_j) - P_j$ by announcing $v_j = \theta_j$, no matter what $v_{-\{i,j\}}$ and $\mathcal{F}$ are. (Notice that it does not even matter whether $\mathcal{F}$ is the true distribution from which θ_{-i} is drawn, because for $\mathcal{M}$ to be classical DST, the players must maximize their utilities by being truthful regardless of the distribution.) Since player j has the same utility in $\mathcal{M}$ and in $\mathbb{M}$, truthfulness in $\mathcal{M}$ implies truthfulness in $\mathbb{M}$ (about his utility type).

Now we prove that, given that all players $j \neq i$ reveal their true utility types, player i maximizes his expected utility by announcing his true knowledge $\mathcal{D}'_i$. Indeed, player i's expected utility from announcing distribution $\mathcal{F}$ is $\mathbb{E}_{v_{-i} \leftarrow \mathcal{D}'_i}[\frac{\epsilon}{2}(2 + \mathtt{BSR}(v_{-i}, \mathcal{F}))]$. Since Brier's scoring rule is strictly proper, this expectation is maximized if and only if $\mathcal{F} = \mathcal{D}'_i$. ■

Lemma 2. *For all assignment contexts C and distributions $\mathcal{D}$,*

$$\text{REV}(\mathbb{M}, C, \mathcal{D}) \geq (1 - \frac{1}{n})\text{OPT}(C, \mathcal{D}) - \epsilon.$$

Proof. Consider the following mental experiment. Fix a valuation profile v drawn from the distribution $\mathcal{D}$, and (if the mechanism is randomized), fix a sequence r of coin flips. Let $\mathcal{M}(v, \mathcal{D}, r)$ be an execution of the optimal mechanism that produces some outcome (A, P). For any player i, let (A^i, P^i) be an outcome such that $A^i_i = \emptyset$, $P^i_i = 0$, $A^i_{-i} = A_{-i}$, and $P^i_{-i} = P_{-i}$. That is, player i gets the empty allocation and pays zero, and all the other players get the same price and allocation as in (A, P). Define the mechanism $\mathcal{M}_i$ which, given a valuation profile v, coin flips r, and distribution $\mathcal{D}$, runs $\mathcal{M}(v, \mathcal{D}, r)$ but gives player i the empty allocation and charges him zero. That is, $\mathcal{M}_i$ produces the outcome $\mathcal{M}_i(v, \mathcal{D}, r) = (A^i, P^i)$. Note that the average revenue of the $\mathcal{M}_i$ mechanisms (taken over the choice of i) is

$$\frac{1}{n}\sum_{i=1}^{n} \text{REV}(\mathcal{M}_i(v, \mathcal{D}, r)) = \frac{1}{n}\sum_{i=1}^{n}\sum_{j \neq i} P_j = \frac{n-1}{n}\sum_{j=1}^{n} P_j = \frac{n-1}{n}\text{REV}(\mathcal{M}(v, \mathcal{D}, r)).$$

Taking expectation over all $v \leftarrow \mathcal{D}$ and all sequences of coin flips r, we get that

$$\frac{1}{n}\sum_{i=1}^{n} \text{REV}(\mathcal{M}_i(\mathcal{D})) = \frac{n-1}{n}\text{REV}(\mathcal{M}(\mathcal{D})) = \frac{n-1}{n}\text{OPT}(C, \mathcal{D}).$$

Since the mechanism $\mathcal{M}$ is classical-DST, and each $\mathcal{M}_i$ does not change the outcomes for any player $j \neq i$, we must have that $\mathcal{M}_i$ is dominant-strategy truthful for all players $j \neq i$. Furthermore, it has the property that it produces an allocation A^i such that $A^i_i = \emptyset$, an allocation that does not serve player i.

Now let's restrain ourselves to outcomes that do not serve player i. Since the mechanism $\mathcal{M}$ is optimal, the maximum expected revenue of classical-DST mechanisms not serving player i is obtained by running $\mathcal{M}$ on inputs $v_{-i}, \mathcal{D}_{-i}$, where v_{-i} is the profile of valuations of all players except i, and $\mathcal{D}_{-i}$ is the distribution of v_{-i} induced by restricting $\mathcal{D}$ to Θ_{-i}. In particular, this generates more expected revenue than running $\mathcal{M}_i$, which is also a classical-DST mechanism that does not serve player i. We can conclude that

$$\text{REV}(\mathcal{M}(\mathcal{D}_{-i})) \geq \text{REV}(\mathcal{M}_i(\mathcal{D})).$$

To continue, we observe that the expected revenue of optimal mechanisms increases with the precision of their type-distributions. That is, modelling them as a single universal mechanism $\mathcal{M}$ receiving the relevant type-distribution as a separate input; letting $\mathcal{P}$ be a partition of the type-space Θ into events; and assuming that, when the true-type profile θ is randomly selected from $\mathcal{D}$, $\mathcal{M}$'s separate input is $\mathcal{D}|E$ —where E is the unique set in $\mathcal{P}$ containing the realized θ— we have

$$\sum_{E\in\mathcal{P}} \Pr_{\theta\leftarrow\mathcal{D}}(\theta\in E)\cdot \textsc{opt}(C,\mathcal{D}|E) \geq \textsc{opt}(C,\mathcal{D}).$$

Accordingly, we must have that the revenue of $\mathcal{M}$ does not decrease when given player i's true knowledge $\mathcal{D}'_i$, which is $\mathcal{D}|S_i$ whenever $\theta \in S_i \in \mathscr{P}_i$. That is,

$$\textsc{rev}(\mathcal{M},-i) \triangleq \textstyle\sum_{S_i\in\mathscr{P}_i} \Pr_{\theta\leftarrow\mathcal{D}}(\theta\in S_i)\cdot \textsc{rev}(\mathcal{M}((\mathcal{D}|S_i)_{-i})) \geq \textsc{rev}(\mathcal{M}(\mathcal{D}_{-i})),$$

where $(\mathcal{D}|S_i)_{-i}$ is the distribution of θ_{-i} induced by restricting $\mathcal{D}|S_i$ to Θ_{-i}.

The mechanism $\mathbb{M}$ generates revenue by choosing a player i at random and running $\mathcal{M}(\theta_{-i},\mathcal{D}'_i)$ —when the players are truthful, $v_{-i}=\theta_{-i}$ and $\mathcal{F}=\mathcal{D}'_i$. Thus the expected revenue (despite the reward in step c) of $\mathbb{M}$ is

$$\tfrac{1}{n}\textstyle\sum_{i=1}^{n} \textsc{rev}(\mathcal{M},-i),$$

which, based on previous inequalities, is greater than or equal to

$$\frac{1}{n}\sum_{i=1}^{n}\textsc{rev}(\mathcal{M}(\mathcal{D}_{-i})) \geq \frac{1}{n}\sum_{i=1}^{n}\textsc{rev}(\mathcal{M}_i(\mathcal{D})) = \frac{n-1}{n}\textsc{opt}(C,\mathcal{D}).$$

Taking into account the reward given in step c, which is at most ϵ, we can conclude that the expected revenue of $\mathbb{M}$ is $\geq \frac{n-1}{n}\textsc{opt}(C,\mathcal{D})-\epsilon$, as desired. ■

Theorem 1 follows directly from Lemmas 1 and 2. *Q.E.D.*

Computational Remarks In our proof of Theorem 1, we use the optimality of mechanism $\mathcal{M}$ only to derive two inequalities: namely inequalities $\textsc{rev}(\mathcal{M},-i)\geq\textsc{rev}(\mathcal{M}(\mathcal{D}_{-i}))$ and $\textsc{rev}(\mathcal{M}(\mathcal{D}_{-i}))\geq\textsc{rev}(\mathcal{M}_i(\mathcal{D}))$. The first can be interpreted as the following condition: "the more precise the type-distribution known to $\mathcal{M}$, the better $\mathcal{M}$'s revenue performance." The second inequality can be interpreted as the following condition: "$\mathcal{M}$ generates more revenue when running with all players but i, than when running on all players and then throwing away i's payment." Thus, if $\mathcal{M}$ satisfies the above two conditions and is approximately optimal, with approximation ratio β, then $\mathbb{M}$ will also be approximately optimal, with approximation ratio $\frac{n-1}{n}\beta$.

When our mechanism $\mathbb{M}$ is explicit, it needs to evaluate Brier's scoring rule. For this, we need to have access to $\mathcal{F}(v_{-i})$ and to compute the norm $\|\mathcal{F}\|_2^2$, which might be computationally expensive when $\mathcal{F}$ is correlated and/or the range of the players' valuations is big. By contrast, a variant of Good's logarithmic scoring rule, $\mathtt{LSR}_{a,b}(v_{-i},\mathcal{F})=a+b\log(\mathcal{F}(v_{-i}))$ with $a,b>0$, is still strictly proper, but only requires access to $\mathcal{F}(v_{-i})$. However, although always providing positive utility in expectation for suitable values of a and b, it has the disadvantage that, for some distributions $\mathcal{F}$ and some valuation subprofiles v_{-i}, it can give arbitrarily negative rewards, giving player i negative utility. If positive utility in expectation is good enough, then $\mathtt{LSR}_{a,b}$ is a perfectly suitable scoring rule.

5 Proof of Theorem 2

The high-level structure of our mechanism $\mathscr{M}$ is very simple. Essentially, the players act only twice.

The first time, they **secretly** transmit their true types to $\mathscr{M}$. There will be enough incentives so that they do so truthfully, and thus $\mathscr{M}$ can publicly announce (1) the *identity* of the (alleged) star player, and (2) the (alleged) true valuations of the other players.

(Note that the above secret transmission makes $\mathscr{M}$ a mechanism of *imperfect information*. Indeed, we do not know how to provide incentives sufficient to guarantee that $\mathscr{M}$ is two-step DST if —say— the players simultaneously but publicly announced their own valuations.)

The second time, each player i not identified as the star player deduces from his original knowledge D'_i, his true type θ_i, and the announcement of $\mathscr{M}$, a more refined distribution about the star player's valuation, which he then announces to $\mathscr{M}$. Again, there will be enough incentives so that this time too the involved players will be truthful. At this point, $\mathscr{M}$ aggregates all the received distributions so as to get a much more refined distribution about the star player's valuation. From this distribution, assuming that this were the only information it had about the star player, $\mathscr{M}$ computes the best take-it-or-leave-it offer to the star player, as if he were the only player around. It also computes another possible offer to the star player: namely the (allegedly) second-highest valuation it previously learned. Finally, it chooses the higher offer, and then decides whether to allocate the good to the star player, and how to charge him, by simulating his acceptance or rejection of the higher offer using his secretly transmitted (alleged) true valuation.

Let us now fix δ and ϵ arbitrarily in $(0, 1/4)$, and provide $\mathscr{M}$'s details. Again, numbered steps are taken by the players, and steps marked by letters are steps taken by the mechanism/auctioneer. Since we are dealing with single-good auctions, we simplify our notation by letting an allocation A be a number in $\{0, 1, \dots, n\}$: $A = 0$ means that the good is unsold, $A = i \neq 0$ means that the good is assigned to player i.

Mechanism $\mathscr{M}$

a. *Set (A, P) to be the empty outcome, that is, $A = 0$ and $P_i = 0$ for each i.*
Comment. (A, P) will be the final outcome of $\mathscr{M}$.

1. *Each player i secretly transmits to the auctioneer an integer $v_i \in [B]$.*
Comment. Allegedly v_i is i's true valuation.

b. *The auctioneer flips a biased coin C_1, such that $C_1 = Heads$ with probability δ.*

c. *If $C_1 = Heads$, the auctioneer does the follows.*

- c_1. *Choose a player $r \in \{1, ..., n\}$ and a price $p \in [B]$ uniformly at random.*
- c_2. *If $v_r \geq p$, then $A = r$ and $P_r = p - \epsilon$.*
- c_3. *The mechanism ends.*

Comment. If $v_r < p$ then the good is unsold. Steps b and c help to ensure that a player is strictly better-off to be truthful about his valuation.

d. *If $C_1 = Tails$, let $* = \operatorname{argmax}_i v_i$. The auctioneer publicly announces $(*, v_{-*})$.*
Comment. Ties are broken lexicographically. Player $*$ is the only candidate for winning the good. The auctioneer informs the players the identity of $*$ and the values announced by the other players.

2. *Each player $j \neq *$ simultaneously announces to the auctioneer a probabilistic distribution $\mathcal{P}^j$ over $[B]$.*
Comment. Allegedly $\mathcal{P}^j$ is j's knowledge about the distribution of player $*$'s true valuation.

e. *The auctioneer computes the following:*

- e_1. $sp = \operatorname{argmax}_{j \neq *} v_j$.
  Comment. Ties are broken lexicographically. Player sp is the "second-valuation" player.
- e_2. *If $sp > *$ then $CP = v_{sp}$, otherwise $CP = v_{sp} + 1$.*
  Comment. CP is the "classical price". Essentially it is the "second price", but because of the way to break ties, the value of CP depends on which one is lexicographically first, player $*$ or player sp.
- e_3. $\widehat{\mathcal{P}} = \mathtt{AGG}(*, (\mathcal{P}^1, \dots, \mathcal{P}^{*-1}, \mathcal{P}^{*+1}, \dots, \mathcal{P}^n))$.
  Comment. Allegedly $\widehat{\mathcal{P}}$ is the aggregated knowledge about the distribution of $*$'s true valuation given the other players' knowledge.
- e_4. $KR = \max_{p \in [B]} \left(p \cdot \Pr_{V \leftarrow \widehat{\mathcal{P}}}[V \geq p]\right)$.
  Comment. KR is the "known revenue", allegedly the maximum expected revenue that can be generated from $*$ given the aggregated knowledge about him.
- e_5. $KP = \operatorname{argmax}_{p \in [B]} \left(p \cdot \Pr_{V \leftarrow \widehat{\mathcal{P}}}[V \geq p]\right)$.
  Comment. KP is the "best known price", allegedly the price that should be used to charge $*$ in order to generate expected revenue KR.

f. *The auctioneer allocates the good and decides $*$'s price with the following rule:*

f_1. *If* $CP \geq KR$ *then* $A = *$ *and* $P_* = CP - \epsilon$.

f_2. *Else, if* $v_* \geq KP$ *then* $A = *$ *and* $P_* = KP - \epsilon$.

Comment. Otherwise (that is, $CP < KR$ and $v_* < KP$), the good is unallocated.

g. *For each* $j \neq *$, $P_j = -\frac{\delta\epsilon}{3nB}(2 + \mathtt{BSR}(v_*, \mathcal{P}^j))$.

Comment. Each player $j \neq *$ is rewarded according to his announced knowledge about $*$.

Analysis of $\mathscr{M}$. The analysis of $\mathscr{M}$ (unlike that of $\mathbb{M}$!) is quite complex and thus given in our Appendix.

Remarks.

- The worst case for $\mathscr{M}$'s revenue is when $\mathcal{D}'_i = \mathcal{D}|\theta_i$. In this case, in fact, the revenue lower bound of Theorem 2 is tight. However, the actual expected revenue generated by $\mathscr{M}$ grows nicely with the quality of the players' knowledge. In particular, when the players "collectively know θ", that is, when the distributions $\mathcal{D}'_i$ are so refined that the intersection of their supports contains a single type profile (necessarily θ), then $\mathscr{M}$'s revenue is arbitrarily close to $\theta_\star$, the maximum revenue ever possible.

  (Again, notice that for the players to collectively know θ, it is not necessary that θ is common knowledge, nor that each player individually knows θ.)

- Mechanism $\mathscr{M}$ trivially accommodates the case when the designer himself has some distributional knowledge about the star player, or the underlying distribution $\mathcal{D}$.

6 Conclusions

By relying on an assumption weaker than traditional Decentralized Assumptions 1, 2, or 3, conservative-Bayesian mechanisms are both more realistic and more challenging. We believe and hope that, given more theoretical attention, they will prove useful in more applications.

Future challenges include dealing with players whose knowledge is approximate and/or represented in algorithmic form, with approximate scoring rules, with continuous valuations, with aggregation of the knowledge of all players, and with broader classes of contexts. We have already started making progress in some of them. Fortunately there is a lot to do!

References

[1] S. Baliga and R. Vohra. Market research and market design. *Advances in Theoretical Economics*, 3(1):1059–1059, 2003.

[2] G.W. Brier. Verification of forecasts expressed in terms of probability. *Monthly weather review*, 78:1–3, 1950.

[3] S. Chawla, J.D. Hartline, D.L. Malec and B. Sivan. Multi-parameter Mechanism Design and Sequential Posted Pricing. In *Proceedings of the 42nd ACM symposium on Theory of computing*, pages 311–320. ACM, 2010.

[4] K.S. Chung and J.C. Ely. Foundations of Dominant-Strategy Mechanisms. *Review of Economic Studies*, 74(2):447–476, 2007.

[5] P. Dhangwatnotai, T. Roughgarden, and Q. Yan. Revenue maximization with a single sample. In *Proceedings of the 11th ACM conference on Electronic commerce*, pages 129–138. ACM, 2010.

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

[7] A.V. Goldberg and J.D. Hartline. Competitiveness via consensus. In *SODA '03: Proceedings of the fourteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 215–222, Philadelphia, PA, USA, 2003. Society for Industrial and Applied Mathematics.

[8] A.V. Goldberg, J.D. Hartline, A.R. Karlin, M. Saks, and A. Wright. Competitive auctions. *Games and Economic Behavior*, 55(2):242–269, 2006.

[9] I.J. Good. Rational Decisions. *Breakthroughs in Statistics: Foundations and basic theory*, page 365, 1992.

[10] J.D. Hartline and T. Roughgarden. Simple versus optimal mechanisms. In *Proceedings of the tenth ACM conference on Electronic commerce*, pages 225–234. ACM, 2009.

[11] R.B. Myerson. Optimal auction design. *Mathematics of operations research*, 6(1):58, 1981.

[12] Z. Neeman. The effectiveness of English auctions. *Games and Economic Behavior*, 43(2):214–238, 2003.

[13] A. Ronen. On approximating optimal auctions. In *Proceedings of the 3rd ACM conference on Electronic Commerce*, pages 11–17. ACM, 2001.

[14] I. Segal. Optimal pricing mechanisms with unknown demand. *The American economic review*, 93(3):509–529, 2003.

[15] R. Wilson. Game-theoretic approaches to trading processes. In *Advances in Economic Theory: Fifth World Congress, ed. by T. Bewley*, pages 33–77, 1987.

Appendix:

Analysis of $\mathscr{M}$

We first provide some convenient notions that will be used throughout the analysis. For any strategy profile γ, and any variable x used in the mechanism $\mathscr{M}$ (e.g., v_i, C_1, r, p, etc), we denote by $x(\gamma)$ the value of this variable in the execution of γ. So $v_i(\gamma)$ is the value announced by player i in Step 1, $C_1(\gamma)$ is the coin flipped in Step b, $r(\gamma)$ and $p(\gamma)$ are the player and the price chosen in Step c_1, in the execution of γ. When the strategy profile γ is clear from the analysis, we may omit it and talk about v_i, C_1, r, p, etc, directly.

We state and prove three lemmas about our mechanism, which together directly imply Theorem 2.

Lemma 3. *It is strictly dominant for each player i to announce his true valuation in Step 1 (i.e., $v_i = \theta_i$).*

Proof. We show that for any true valuation profile θ and any pure strategy σ_i which announces $v_i \neq \theta_i$ in Step 1, there exists a pure strategy $\widehat{\sigma}_i$ which strictly dominates σ_i, that is, for all pure strategy subprofiles τ_{-i},

$$\mathbb{E}[u_i(\widehat{\sigma}_i \sqcup \tau_{-i})|\theta] - \mathbb{E}[u_i(\sigma_i \sqcup \tau_{-i})|\theta] > 0,$$

where the expectations are taken over the coins tossed by $\mathscr{M}$. Since the true valuation profile θ is always fixed, and every event and every expectation are conditioned on θ, we omit the conditioning on θ in the analysis below, for a more succinct presentation.

Strategy $\widehat{\sigma}_i$ works as follows. In Step 1, player i announces θ_i. In Step 2 (whenever reached by the mechanism and whenever $* \neq i$), denoting the valuations announced by the auctioneer about the other players as $v_{-\{*,i\}}$, player i announces $\mathcal{P}^i$ to be the same as what σ_i would have announced given the announcement of the auctioneer in Step d being $(*, v_i \sqcup v_{-\{*,i\}})$ (according to σ_i, player i announced v_i instead of θ_i in Step 1).

Let $\sigma = \sigma_i \sqcup \tau_{-i}$ and $\widehat{\sigma} = \widehat{\sigma}_i \sqcup \tau_{-i}$. Notice that for any strategy profile γ, the expected utility of player i is

$$\begin{aligned}\mathbb{E}[u_i(\gamma)] &= \Pr[C_1(\gamma) = Heads] \cdot \mathbb{E}[u_i(\gamma)|C_1(\gamma) = Heads] + \Pr[C_1(\gamma) = Tails] \cdot \mathbb{E}[u_i(\gamma)|C_1(\gamma) = Tails] \\ &= \frac{\delta}{nB} \cdot \sum_{k=0}^{v_i(\gamma)} (\theta_i - k + \epsilon) + (1-\delta) \cdot \mathbb{E}[u_i(\gamma)|C_1(\gamma) = Tails]. \qquad (1)\end{aligned}$$

According to Equation 1, we have that

$$\begin{aligned}&\mathbb{E}[u_i(\widehat{\sigma})] - \mathbb{E}[u_i(\sigma)] \\ = &\left(\frac{\delta}{nB} \cdot \sum_{k=0}^{v_i(\widehat{\sigma})} (\theta_i - k + \epsilon) + (1-\delta) \cdot \mathbb{E}[u_i(\widehat{\sigma})|C_1(\widehat{\sigma}) = Tails]\right) \\ &- \left(\frac{\delta}{nB} \cdot \sum_{k=0}^{v_i(\sigma)} (\theta_i - k + \epsilon) + (1-\delta) \cdot \mathbb{E}[u_i(\sigma)|C_1(\sigma) = Tails]\right) \\ = &\frac{\delta}{nB} \cdot \left(\sum_{k=0}^{v_i(\widehat{\sigma})} (\theta_i - k + \epsilon) - \sum_{k=0}^{v_i(\sigma)} (\theta_i - k + \epsilon)\right) \\ &+ (1-\delta) \cdot \left(\mathbb{E}[u_i(\widehat{\sigma})|C_1(\widehat{\sigma}) = Tails] - \mathbb{E}[u_i(\sigma)|C_1(\sigma) = Tails]\right).\end{aligned}$$

Thus to prove $\mathbb{E}[u_i(\widehat{\sigma})] - \mathbb{E}[u_i(\sigma)] > 0$, it suffices to prove that

$$\sum_{k=0}^{v_i(\widehat{\sigma})} (\theta_i - k + \epsilon) - \sum_{k=0}^{v_i(\sigma)} (\theta_i - k + \epsilon) \geq \epsilon, \qquad (2)$$

and that

$$\mathbb{E}[u_i(\widehat{\sigma})|C_1(\widehat{\sigma}) = Tails] - \mathbb{E}[u_i(\sigma)|C_1(\sigma) = Tails] > -\frac{\delta\epsilon}{nB}. \tag{3}$$

To prove Equation 2, notice that $v_i(\widehat{\sigma}) = \theta_i$ by definition of $\widehat{\sigma}_i$. If $v_i(\sigma) < \theta_i$, then the left-hand side of Equation 2 becomes

$$\sum_{k=v_i(\sigma)+1}^{\theta_i} (\theta_i - k + \epsilon) \geq \theta_i - \theta_i + \epsilon = \epsilon.$$

If $v_i(\sigma) > \theta_i$, then the left-hand side of Equation 2 becomes

$$\sum_{k=\theta_i+1}^{v_i(\sigma)} (k - \theta_i - \epsilon) \geq \theta_i + 1 - \theta_i - \epsilon = 1 - \epsilon > \epsilon,$$

where the last inequality is because $\epsilon \in (0, 1/4)$. Thus Equation 2 holds.

To prove Equation 3, we distinguish four cases.

Case 1. $*(\widehat{\sigma}) \neq i$ and $*(\sigma) \neq i$.

In this case, since for any other player j, j's strategy is τ_j in both $\widehat{\sigma}$ and σ, we have that $v_j(\widehat{\sigma}) = v_j(\sigma)$, and thus $*(\widehat{\sigma}) = *(\sigma)$. Therefore $\mathcal{P}^i(\widehat{\sigma}) = \mathcal{P}^i(\sigma)$, by definition of $\widehat{\sigma}$. Accordingly, the reward player i receives in Step g is the same in both $\widehat{\sigma}$ and σ, that is, $P_i(\widehat{\sigma}) = P_i(\sigma)$, implying that the left-hand side of Equation 3 is precisely 0, and that Equation 3 holds.

Case 2. $*(\widehat{\sigma}) = *(\sigma) = i$.

In this case, for any other player j, $v_j(\widehat{\sigma}) = v_j(\sigma)$, and thus the announcements of the auctioneer in Step d are the same in both $\widehat{\sigma}$ and σ. Therefore we have that:

$\mathcal{P}^j(\widehat{\sigma}) = \mathcal{P}^j(\sigma)$,
$sp(\widehat{\sigma}) = sp(\sigma)$,
$CP(\widehat{\sigma}) = CP(\sigma) \leq \theta_i$ (no matter whether the "second-valuation" player is before or after player i),
$KR(\widehat{\sigma}) = KR(\sigma)$, and
$KP(\widehat{\sigma}) = KP(\sigma)$.

Accordingly, when $CP(\sigma) \geq KR(\sigma)$, or when $CP(\sigma) < KR(\sigma)$ and $KP(\sigma) > \max\{\theta_i, v_i(\sigma)\}$, or when $CP(\sigma) < KR(\sigma)$ and $KP(\sigma) \leq \min\{\theta_i, v_i(\sigma)\}$, the expected utilities of player i in the execution of $\widehat{\sigma}$ and in the execution of σ are the same, and thus the left-hand side of Equation 3 is precisely 0.

When $CP(\sigma) < KR(\sigma)$ and $\min\{\theta_i, v_i(\sigma)\} < KP(\sigma) \leq \max\{\theta_i, v_i(\sigma)\}$, player i's expected utilities are different in the execution of $\widehat{\sigma}$ and in the execution of σ. Indeed, in this situation, if $\theta_i > v_i(\sigma)$, then i gets the good and pays a price which is at least ϵ less than his true valuation according to $\widehat{\sigma}$, while gets nothing and pays nothing according to σ. Thus we have that

$$\mathbb{E}[u_i(\widehat{\sigma})|C_1(\widehat{\sigma}) = Tails] \geq \epsilon \quad \text{and} \quad \mathbb{E}[u_i(\sigma)|C_1(\sigma) = Tails] = 0.$$

If $\theta_i < v_i(\sigma)$, then i gets nothing and pays nothing according to $\widehat{\sigma}$, while gets the good and pays at least $1 - \epsilon$ more than his true valuation according to σ. Thus we have that

$$\mathbb{E}[u_i(\widehat{\sigma})|C_1(\widehat{\sigma}) = Tails] = 0 \quad \text{and} \quad \mathbb{E}[u_i(\sigma)|C_1(\sigma) = Tails] < -1 + \epsilon.$$

Thus Equation 3 always holds.

Case 3. $*(\widehat{\sigma}) = i$ and $*(\sigma) \neq i$.

In this case, similar to Case 2, in the execution of $\widehat{\sigma}$, no matter what the other players do in Step 2, player i never pays more than his true valuation when he gets the good — indeed, $CP(\widehat{\sigma}) \leq v_i(\widehat{\sigma}) = \theta_i$ due to the way to break ties and the way to set the value of CP; and i pays $KP(\widehat{\sigma}) - \epsilon$ only if $KP(\widehat{\sigma}) \leq v_i(\widehat{\sigma})$. Thus we have that

$$\mathbb{E}[u_i(\widehat{\sigma})|C_1(\widehat{\sigma}) = Tails] \geq 0.$$

While in the execution of σ, the only utility player i gets is the reward he receives in Step g, which is always less than $\frac{\delta\epsilon}{nB}$ since the Brier scoring rule is always in $[-2, 0]$, implying that

$$\mathbb{E}[u_i(\sigma)|C_1(\sigma) = Tails] = -P_i(\sigma) < \frac{\delta\epsilon}{nB}.$$

Thus Equation 3 holds.

Case 4. $*(\widehat{\sigma}) \neq i$ and $*(\sigma) = i$.

This case implies that $v_i(\sigma) > \theta_i$. Further, it implies that $CP(\sigma) > \theta_i$ — indeed, if $sp(\sigma) < i$ then $\theta_i \leq v_{sp(\sigma)}(\sigma) = CP(\sigma) - 1$, while if $sp(\sigma) > i$ then $\theta_i < v_{sp(\sigma)}(\sigma) = CP(\sigma)$. Because both θ_i and $CP(\sigma)$ are integers, we have that $CP(\sigma) \geq \theta_i + 1$. Accordingly, when σ is played, whenever player i gets the good, his price is at least $CP(\sigma) - \epsilon$, which is at least $1 - \epsilon$ more than his true valuation. Thus we have that

$$\mathbb{E}[u_i(\sigma)|C_1(\sigma) = Tails] \leq 0.$$

While when $\widehat{\sigma}$ is played, player i pays nothing but gets some small reward in Step g, thus we have that

$$\mathbb{E}[u_i(\widehat{\sigma})|C_1(\widehat{\sigma}) = Tails] > 0.$$

Therefore Equation 3 holds.

In sum, Equations 2 and 3 both hold, implying that $\mathbb{E}[u_i(\widehat{\sigma})] - \mathbb{E}[u_i(\sigma)] > 0$. ■

Lemma 4. *If all players are truthful in Step 1, then in Step 2 it is strictly dominant for each player $i \neq *$ to truthfully report his knowledge about $*$. That is, to report $\mathcal{P}^i = \mathcal{D}'_i|(*, v_{-*})$, $\mathcal{D}'_i$ conditioned on $* = \star$ and $\theta_{-*} = v_{-*}$.*[3]

Proof. Let Σ^t be the profile of strategy sets such that for each player j, Σ^t_j consists of all strategies of player j that always announce θ_j in Step 1, no matter what the value of θ_j is (recall that in a Bayesian setting, a strategy of a player j must specify what player j does for any possible value of θ_j). Consider an arbitrary pure strategy $\sigma_i \in \Sigma^t_i$ such that there exists a player $k \neq i$ and valuation subprofile $w_{-k} \in [B]^{n-1}$ satisfying the following: (1) $w_i = \theta_i$; (2) $\Pr_{\theta \leftarrow \mathcal{D}'_i}[(\star, \theta_{-\star}) = (k, w_{-k})] > 0$; and (3) σ_i announces $\mathcal{P}^i \neq \mathcal{D}'_i|(k, w_{-k})$ in Step 2 when the announcement of the auctioneer in Step d is (k, w_{-k}). We prove that there exists another pure strategy $\widehat{\sigma}_i \in \Sigma^t_i$ which strictly dominates σ_i over Σ^t, that is, for any pure strategy subprofile $\tau_{-i} \in \Sigma^t_{-i}$,

$$\mathbb{E}[u_i(\widehat{\sigma}_i \sqcup \tau_{-i})] > \mathbb{E}[u_i(\sigma_i \sqcup \tau_{-i})],$$

where the expectations are taken over the choice of θ according to $\mathcal{D}'_i$, and the coins used by $\mathscr{M}$. Notice that differently from the analysis of Lemma 3, here the expectation really depends on the distribution of θ from i's point of view, that is, depends on $\mathcal{D}'_i$.

Strategy $\widehat{\sigma}_i$ works as follows. In Step 1, player i announces θ_i. In Step 2 (when reached), if the announcement in Step d is not (k, w_{-k}) then i announces the same as σ_i would have announced in this case; otherwise i announces $\mathcal{D}'_i|(k, w_{-k})$.

Let $\sigma = (\sigma_i \sqcup \tau_{-i})$ and $\widehat{\sigma} = (\widehat{\sigma}_i \sqcup \tau_{-i})$. Notice that for any strategy profile γ, player i's expected utility is

$$\begin{aligned}\mathbb{E}[u_i(\gamma)] = & \Pr_{\theta \leftarrow \mathcal{D}'_i}[(\star, \theta_{-\star}) \neq (k, w_{-k})]\mathbb{E}[u_i(\gamma)|(\star, \theta_{-\star}) \neq (k, w_{-k})] \\ & + \Pr_{\theta \leftarrow \mathcal{D}'_i}[(\star, \theta_{-\star}) = (k, w_{-k})]\mathbb{E}[u_i(\gamma)|(\star, \theta_{-\star}) = (k, w_{-k})] \end{aligned} \quad (4)$$

When $(\star, \theta_{-\star}) \neq (k, w_{-k})$, the execution of $\widehat{\sigma}$ and that of σ coincide with each other — in particular, every player announces his true valuation in Step 1, the auctioneer announces $(\star, \theta_{-\star})$ in Step d (when reached), and player i announces $\mathcal{P}^i(\widehat{\sigma}) = \mathcal{P}^i(\sigma)$ in Step 2. Thus

$$\mathbb{E}[u_i(\widehat{\sigma})|(\star, \theta_{-\star}) \neq (k, w_{-k})] = \mathbb{E}[u_i(\sigma)|(\star, \theta_{-\star}) \neq (k, w_{-k})],$$

[3]Recall that given the true valuation profile θ, $\star = \mathrm{argmax}_i\, \theta_i$ with ties broken lexicographically.

which together with Equation 4 implies

$$\begin{aligned}&\mathbb{E}[u_i(\widehat{\sigma})] - \mathbb{E}[u_i(\sigma)]\\ = \ &\Pr_{\theta\leftarrow\mathcal{D}'_i}[(\star,\theta_{-\star})=(k,w_{-k})]\left\{\mathbb{E}[u_i(\widehat{\sigma})|(\star,\theta_{-\star})=(k,w_{-k})] - \mathbb{E}[u_i(\sigma)|(\star,\theta_{-\star})=(k,w_{-k})]\right\}.\end{aligned}$$

Because $\Pr_{\theta\leftarrow\mathcal{D}'_i}[(\star,\theta_{-\star})=(k,w_{-k})]>0$ by hypothesis, to prove $\mathbb{E}[u_i(\widehat{\sigma})] > \mathbb{E}[u_i(\sigma)]$, it suffices to prove

$$\mathbb{E}[u_i(\widehat{\sigma})|(\star,\theta_{-\star})=(k,w_{-k})] > \mathbb{E}[u_i(\sigma)|(\star,\theta_{-\star})=(k,w_{-k})].$$

Further notice that player i announces θ_i in Step 1 in both $\widehat{\sigma}$ and σ, thus his expected utility when $C_1 = Heads$ is the same, that is,

$$\mathbb{E}[u_i(\widehat{\sigma})|(\star,\theta_{-\star})=(k,w_{-k}), C_1(\widehat{\sigma})=Heads] = \mathbb{E}[u_i(\sigma)|(\star,\theta_{-\star})=(k,w_{-k}), C_1(\sigma)=Heads].$$

Since $\Pr[C_1(\widehat{\sigma})=Heads] = \Pr[C_1(\sigma)=Heads]=\delta$, it suffices to prove that

$$\mathbb{E}[u_i(\widehat{\sigma})|(\star,\theta_{-\star})=(k,w_{-k}), C_1(\widehat{\sigma})=Tails] > \mathbb{E}[u_i(\sigma)|(\star,\theta_{-\star})=(k,w_{-k}), C_1(\sigma)=Tails]. \tag{5}$$

To prove Equation 5, notice that in both $\widehat{\sigma}$ and σ, player i's utility solely comes from the reward he receives in Step g, which solely depends on the values of v_* and $\mathcal{P}^i$. Thus it suffices to prove that

$$\begin{aligned}&\mathbb{E}\left[\frac{\delta\epsilon}{nB}(2+\mathtt{BSR}(v_*(\widehat{\sigma}),\mathcal{P}^i(\widehat{\sigma})))\middle|(\star,\theta_{-\star})=(k,w_{-k}), C_1(\widehat{\sigma})=Tails\right]\\ > \ &\mathbb{E}\left[\frac{\delta\epsilon}{nB}(2+\mathtt{BSR}(v_*(\sigma),\mathcal{P}^i(\sigma)))\middle|(\star,\theta_{-\star})=(k,w_{-k}), C_1(\sigma)=Tails\right],\end{aligned}$$

that is, to prove that

$$\begin{aligned}&\mathbb{E}\left[\mathtt{BSR}(v_*(\widehat{\sigma}),\mathcal{P}^i(\widehat{\sigma}))|(\star,\theta_{-\star})=(k,w_{-k}), C_1(\widehat{\sigma})=Tails\right]\\ > \ &\mathbb{E}\left[\mathtt{BSR}(v_*(\sigma),\mathcal{P}^i(\sigma))|(\star,\theta_{-\star})=(k,w_{-k}), C_1(\sigma)=Tails\right].\end{aligned}\tag{6}$$

Notice that in both executions, $*=\star=k$, $v_{-*}=\theta_{-\star}=w_{-k}$, and $v_*=\theta_\star$. Therefore the fact that the auctioneer announces $(*,v_{-*})$ in Step d is equivalent to the fact that player i is informed with $(\star,\theta_{-\star})=(k,w_{-k})$. Thus $\mathcal{P}^i(\widehat{\sigma})=\mathcal{D}'_i|(*,v_{-*})=\mathcal{D}'_i|(k,w_{-k})=\mathcal{D}'_i|(\star,\theta_{-\star})$, which is the true distribution of $\theta_\star$, and thus of $v_*(\widehat{\sigma})$ and $v_*(\sigma)$, from player i's point of view. Accordingly, to prove Equation 6 it suffices to prove that

$$\mathbb{E}_{\theta_\star\leftarrow\mathcal{D}'_i|(\star,\theta_{-\star})}[\mathtt{BSR}(\theta_\star,\mathcal{D}'_i|(\star,\theta_{-\star}))] > \mathbb{E}_{\theta_\star\leftarrow\mathcal{D}'_i|(\star,\theta_{-\star})}[\mathtt{BSR}(\theta_\star,\mathcal{P}^i(\sigma))]. \tag{7}$$

Notice that the coins used by the mechanism have been removed, since the distribution of true valuations does not depend on the mechanism.

While Equation 7 follows directly from the fact that the Brier scoring rule is strictly proper. Therefore $\mathbb{E}[u_i(\widehat{\sigma})] > \mathbb{E}[u_i(\sigma)]$, and Lemma 4 holds. ■

Lemma 5. *When the players are truthful in both Step 1 and Step 2, $\mathscr{M}$'s expected revenue is greater than or equal to $(1-\delta)\cdot(\mathscr{S}-2\epsilon)$.*

Proof. Let τ be the truthful strategy profile, and $\mathbb{E}[\text{REV}(\mathscr{M}(\tau))]$ be the expected revenue generated by $\mathscr{M}$ under τ, where the expectation is taken over $\mathcal{D}$ and the coins used by $\mathscr{M}$. When $C_1 = Heads$, which happens with probability δ, the (expected) revenue that $\mathscr{M}$ gets is non-negative. Therefore it suffices to prove that when $C_1=Tails$, the expected revenue that $\mathscr{M}$ gets is at least $\mathscr{S}-2\epsilon$, that is, to prove that

$$\mathbb{E}[\text{REV}(\mathscr{M}(\tau))|C_1=Tails] \geq \mathscr{S}-2\epsilon.$$

Because

$$\mathbb{E}[\text{REV}(\mathscr{M}(\tau))|C_1 = Tails] = \sum_{i=1}^{n} \sum_{t\in[B]^{n-1}} \Pr_{\theta\leftarrow\mathcal{D}}(\star = i, \theta_{-\star} = t) \cdot \mathbb{E}[\text{REV}(\mathscr{M}(\tau))|C_1 = Tails, \star = i, \theta_{-\star} = t],$$

and

$$\mathscr{S} = \sum_{i=1}^{n} \sum_{t\in[B]^{n-1}} \Pr_{\theta\leftarrow\mathcal{D}}(\star = i, \theta_{-\star} = t) \max_p \left(p \cdot \Pr_{\theta\leftarrow\mathcal{D}}(\theta_\star \geq p|\star = i, \theta_{-\star} = t_{-i})\right),$$

it suffices to prove that for any player i and any $t \in [B]^{n-1}$ such that $\Pr_{\theta\leftarrow\mathcal{D}}(\star = i, \theta_{-\star} = t) > 0$,

$$\mathbb{E}[\text{REV}(\mathscr{M}(\tau))|C_1 = Tails, \star = i, \theta_{-\star} = t] \geq \max_p \left(p \cdot \Pr_{\theta\leftarrow\mathcal{D}}(\theta_\star \geq p|\star = i, \theta_{-\star} = t_{-i})\right) - 2\epsilon. \tag{8}$$

Letting $S(i,t) = \{\theta \in \Theta : \star = i, \theta_{-\star} = t\}$, and letting $\mathcal{D}(i,t)$ be the marginal distribution of θ_i according to $\mathcal{D}|S(i,t)$, the right-hand side of Equation 8 becomes

$$\max_p \left(p \cdot \Pr_{\theta_i\leftarrow\mathcal{D}(i,t)}(\theta_i \geq p)\right) - 2\epsilon.$$

Letting $p(i,t) = \text{argmax}_p \left(p \cdot \Pr_{\theta_i\leftarrow\mathcal{D}(i,t)}(\theta_i \geq p)\right)$, which depends only on (i,t) but not the precise value of θ_i, the above formula further becomes

$$p(i,t) \cdot \Pr_{\theta_i\leftarrow\mathcal{D}(i,t)}(\theta_i \geq p(i,t)) - 2\epsilon. \tag{9}$$

Below we derive another formula for the left-hand side of Equation 8, so that the two sides can be compared easily. Notice that conditioned on $\star = i$ and $\theta_{-\star} = t$, the announcement of the auctioneer in Step d in the execution of τ is (i,t), and thus each player $j \neq i$ is informed that $\star = i$ and $\theta_{-\star} = t$. Accordingly, player j's knowledge about θ, and in particular about θ_i, is further refined.

Indeed, letting $M(i,t)$ be $S(i,t)$ projected on the i-th component, there exists a partition $\mathscr{P}_j(i,t)$ of $M(i,t)$ such that $\mathscr{P}_j(i,t)$ is obtained by first intersecting $S(i,t)$ with each set in $\mathscr{P}_j$ and then projecting on the i-th component. When θ_i is randomly selected from $\mathcal{D}(i,t)$, letting S_j be the unique set in $\mathscr{P}_j$ containing $\theta_i \sqcup t$, player j's knowledge about θ is

$$\mathcal{D}'_j|S(i,t) = (\mathcal{D}|S_j)|S(i,t) = \mathcal{D}|(S_j \cap S(i,t)) = (\mathcal{D}|S(i,t))|(S_j \cap S(i,t)).$$

Therefore j's knowledge about θ_i is $\mathcal{D}(i,t)|S_j(i,t)$, where $S_j(i,t)$ is $S_j \cap S(i,t)$ projected on the i-th component, which is the unique set in $\mathscr{P}_j(i,t)$ containing θ_i. Because player j is truthful in Step 2 in the execution of τ, we have that

$$\mathcal{P}^j = \mathcal{D}(i,t)|S_j(i,t).$$

By definition of the aggregator `AGG`, we have that in the execution of τ,

$$\widehat{\mathcal{P}} = \mathcal{D}(i,t)| \cap_{j\neq i} S_j(i,t).$$

Indeed, there exists another partition $\mathscr{P}(i,t)$ of $M(i,t)$ such that: (1) $\mathscr{P}(i,t)$ is obtained by intersection all $\mathscr{P}_j(i,t)$'s together, and (2) when θ_i is randomly selected from $\mathcal{D}(i,t)$,

$$\widehat{\mathcal{P}} = \mathcal{D}(i,t)|E(i,t),$$

where $E(i,t) = \cap_{j\neq i} S_j(i,t)$ with each $S_j(i,t)$ being the unique set in $\mathscr{P}_j(i,t)$ containing θ_i, and thus $E(i,t)$ is the unique set in $\mathscr{P}(i,t)$ containing θ_i.

With respect to $\mathscr{P}(i,t)$, the left-hand side of Equation 8 can be rewritten as

$$\sum_{E(i,t)\in\mathscr{P}(i,t)} \Pr_{\theta_i\leftarrow\mathcal{D}(i,t)}(\theta_i \in E(i,t)) \cdot \mathbb{E}_{\theta_i\leftarrow\mathcal{D}(i,t)}[\text{REV}(\mathscr{M}(\tau))|C_1 = Tails, \theta_i \in E(i,t), \theta_{-i} = t],$$

which is equal to

$$\sum_{E(i,t)\in\mathscr{P}(i,t)} \Pr_{\theta_i\leftarrow\mathcal{D}(i,t)}(\theta_i \in E(i,t)) \cdot \mathbb{E}_{\theta_i\leftarrow\mathcal{D}(i,t)|E(i,t)}[\text{REV}(\mathscr{M}(\tau))|C_1 = Tails, \theta_i, \theta_{-i} = t]. \tag{10}$$

Notice that the expectation is only taken over the distribution of θ_i, since conditioned on $C_1 = Tails$, the mechanism is deterministic.

Recall that the right-hand side of Equation 8 is equal to Equation 9, which can be rewritten as

$$\sum_{E(i,t)\in\mathscr{P}(i,t)} \Pr_{\theta_i\leftarrow\mathcal{D}(i,t)}(\theta_i \in E(i,t)) \cdot \left(p(i,t) \cdot \Pr_{\theta_i\leftarrow\mathcal{D}(i,t)|E(i,t)}(\theta_i \geq p(i,t)) - 2\epsilon\right).$$

Therefore combining with Equation 10 we have that, to prove Equation 8, it suffices to prove that for each $E(i,t) \in \mathscr{P}(i,t)$ such that $\Pr_{\theta_i\leftarrow\mathcal{D}(i,t)}(\theta_i \in E(i,t)) > 0$,

$$\mathbb{E}_{\theta_i\leftarrow\mathcal{D}(i,t)|E(i,t)}[\text{REV}(\mathscr{M}(\tau))|C_1 = Tails, \theta_i, \theta_{-i} = t] \geq p(i,t) \cdot \Pr_{\theta_i\leftarrow\mathcal{D}(i,t)|E(i,t)}(\theta_i \geq p(i,t)) - 2\epsilon. \tag{11}$$

To prove Equation 11, notice that when θ_i is selected from $\mathcal{D}(i,t)|E(i,t)$, we always have $\widehat{\mathcal{P}} = \mathcal{D}(i,t)|E(i,t)$, and thus

$$KR = \max_p \left(p \cdot \Pr_{\theta_i\leftarrow\mathcal{D}(i,t)|E(i,t)}(\theta_i \geq p)\right) \geq\ p(i,t) \cdot \Pr_{\theta_i\leftarrow\mathcal{D}(i,t)|E(i,t)}(\theta_i \geq p(i,t)). \tag{12}$$

Therefore by solely selling the good to i and charging him KP whenever $KP \leq v_i(= \theta_i)$, the mechanism would have already generated expected revenue $\geq KR - 2\epsilon$, where one ϵ is the discount given to i, and the other ϵ is an upper bound on the total reward given to other players in Step g.

Further notice that conditioned on $\star = i$ and $\theta_{-\star} = t$, the second-valuation player sp and thus the classical price CP are totally determined. Therefore the comparison between CP and KR in Step f does not alter the distribution from which θ_i is drawn, that is, $\mathcal{D}(i,t)|E(i,t)$. In fact, CP is the lowest value of θ_i for i to be the star player given $\theta_{-i} = t$: $CP = \theta_{sp}$ if $sp > i$, and $CP = \theta_{sp} + 1$ if $sp < i$. Accordingly, for any θ_i in the support of $\mathcal{D}(i,t)$ and thus in the support of $\mathcal{D}(i,t)|E(i,t)$, we have that $\theta_i \geq CP$, which implies that

$$KR \geq CP \cdot \Pr_{\theta_i\leftarrow\mathcal{D}(i,t)|E(i,t)}(\theta_i \geq CP) = CP \cdot 1 = CP.$$

Thus even by comparing CP with KR and selling the good to i accordingly, the mechanism still generates expected revenue $\geq KR - 2\epsilon$. Indeed, if $CP = KR$ then with probability 1 the revenue generated is at least $CP - 2\epsilon = KR - 2\epsilon$, otherwise the expected revenue generated is at least

$$(KP - \epsilon) \cdot \Pr_{\theta_i\leftarrow\mathcal{D}(i,t)|E(i,t)}(\theta_i \geq KP) - \epsilon \geq KR - 2\epsilon.$$

Accordingly,

$$\mathbb{E}_{\theta_i\leftarrow\mathcal{D}(i,t)|E(i,t)}[\text{REV}(\mathscr{M}(\tau))|C_1 = Tails, \theta_i, \theta_{-i} = t] \geq KR - 2\epsilon,$$

which together with Equation 12 implies Equation 11, and thus the revenue lower bound in Lemma 5. ■

At this point Theorem 2 can be proven directly. First of all, Lemmas 3 and 4 imply that $\mathscr{M}$ is two-step DST. Second, although the revenue lower bound to be proven according to Theorem 2 is $(1-\delta)\mathscr{S} - \epsilon$, and given parameters δ and ϵ the expected revenue of $\mathscr{M}$ according to Lemma 5 is $\geq (1-\delta)(\mathscr{S} - 2\epsilon)$, such a mismatch can be easily solved, because the parameters of $\mathscr{M}$ can be chosen to be arbitrarily small. More precisely, taking $\delta' < \delta$ and $\epsilon' < \frac{\epsilon}{2(1-\delta')}$ and running $\mathscr{M}$ with parameters δ' and ϵ', the expected revenue generated is at least $(1-\delta')(\mathscr{S} - 2\epsilon') = (1-\delta')\mathscr{S} - 2(1-\delta')\epsilon' > (1-\delta)\mathscr{S} - \epsilon$, as desired. *Q.E.D.*