Guaranteeing Perfect Revenue From Perfectly Informed Players

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

Avinatan Hassidim
RLE, MIT
Cambridge, MA 02139, USA
avinatanh@gmail.com

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

September 19, 2009

Abstract

Maximizing revenue in the presence of perfectly informed players is a well known goal in mechanism design. Yet, all current mechanisms for this goal are vulnerable to equilibrium selection and therefore far from guaranteeing that maximum revenue will be obtained. In this paper we both clarify and rectify this situation by

- Proving that no *weakly dominant-strategy* mechanism (traditionally considered immune to equilibrium selection) guarantees an arbitrarily small fraction of the maximum possible revenue;

and, more importantly,

- Constructing a *robust-strategy* mechanism (a new type of mechanism provably immune to equilibrium selection) guaranteeing a fraction arbitrarily close to 1 of the maximum possible revenue.

In particular, therefore, we provably separate implementation in dominant-strategies from implementation in robust strategies. Our robust-strategy mechanism actually is of a stronger type. Namely, it is of extensive-form and has a *unique* sub-game-perfect equilibrium.

In addition, our mechanism guarantees the players' maximum privacy and withstands rational collusion to the largest possible extent. This is important, since both privacy and collusion are typically capable of derailing the intended functioning of a mechanism.

1 Introduction

1.1 Classical Mechanism Design

Contexts and mechanisms. A *context* C describes the players, the outcomes and the players' preferences over the outcomes. A *mechanism* M describes the *strategies* available to the players, and how strategies determine outcomes. Together, a context C and a mechanism M define a *game* G, $G = (C, M)$, in which each rational player will endeavor to choose his own strategy so as to maximize his own utility.

Mechanism design. Mechanism design aims at finding a mechanism M such that, for any context C (or any C in a given class), a desired property $\mathbb{P}$ holds for the outcomes of the game (C, M), when rationally played. The difficulty is that the designer does not exactly know the players' preferences, while $\mathbb{P}$ typically depends on such preferences. In the purest form of mechanism design, *all knowledge about the players lies with the players themselves.* The designer can count only on the players' rationality. And based solely on this fact, he must design M so that it becomes "in the best interest of the players" to satisfy $\mathbb{P}$. That is, he must ensure that $\mathbb{P}$ holds in a rational play of M. But: *What is a rational play?*

The classical interpretation of a rational play. The classical interpretation of a rational play is an *equilibrium*, that is a profile of strategies $\sigma = \sigma_1, \ldots, \sigma_n$ such that no player i has an incentive to deviate from his specified strategy σ_i to an alternative strategy σ_i'. But equilibria are vastly different in their "quality." The weakest form is that of a *Nash equilibrium,* simply stating that i prefers σ_i to any alternative σ_i' only if he *believes that every other player j will stick to his specified σ_j*. That is, Nash equilibrium only guarantees that i prefers $\sigma_1, \ldots, \sigma_i, \ldots, \sigma_n$ to $\sigma_1, \ldots, \sigma_i', \ldots, \sigma_n$. If σ is a *dominant-strategy equilibrium,* the strongest form of equilibrium, then, for any player i, σ_i is i's best strategy no matter what strategies the other players may choose. More precisely, a dominant-strategy equilibrium σ is called strict (respectively, weak) if, for any player i, any alternative strategy σ_i', and any strategy sub-profile τ_{-i} for the other players, i's utility is strictly larger (respectively, larger or equal to) when i plays σ_i than when he plays σ_i'.

1.2 Our Goal

This paper focuses on a classical context: *quasi-linear utilities with non-negative valuations.* Namely, there are finitely many possible *states*, $\omega_1, \ldots, \omega_k$, including the *null state*, which every player values 0; each player i has non-negative value $v_i(\omega_j)$ for each state ω_j; each outcome consists of a state ω together with a price P_i for each player i; and the utility of each player i for such an outcome is $v_i(\omega) - P_i$. (The *revenue* of an outcome (ω, P) consists of $\sum_i P_i$. The function v_i is i's *valuation.*)

Such context models a great deal of situations. For instance, in an auction of multiple goods, a state ω represents which player wins which items. Accordingly, the utility of player i in an outcome (ω, P) naturally is his value for the items he gets in ω, minus the price he pays. In another example, each state ω represents one of finitely many ways of building a bridge across a given river. Accordingly, and naturally too, each player has different values for each possible bridge. (For instance, a player's value for a given potential bridge may dependent on how distant it would be from his house.) The list of examples could go on and on. In all of them, however, no matter what the mechanisms may be, it is also natural for different subsets of the players to collude —that is, to coordinate their strategies— so as to improve their utilities.

In such a classical context, our goal is equally classical: getting an outcome of maximum revenue when the players have *perfect knowledge.* That is, when each player knows the valuations of all players (as well as who colludes with whom, if collusion exists among the players).

When the players' knowledge is best possible, it is natural to ask whether the best possible revenue can be obtained. Note that, without the ability of imposing arbitrary prices, the best possible revenue that a mechanism can hope to get from rational players is the *maximum social welfare*, that is, $\max_\omega \sum_i v_i(\omega)$. Thus:

Can a mechanism guarantee perfect revenue from perfectly informed players?

1.3 Three Main Obstacles

Equilibrium Selection Plenty of mechanisms have been proposed for our goal. Yet, *none of them achieves it in a robust way.* A main obstacle on their way is *equilibrium selection.* Let us explain.

It should be realized that designing a mechanism so as to guarantee a property $\mathbb{P}$ "at a Nash Equilibrium" is a weak guarantee. First, because there may be several Nash equilibria, while $\mathbb{P}$ holds for just some of them. Moreover, even if $\mathbb{P}$ held for all equilibria, $\mathbb{P}$ may not hold at all in a real play. For instance, assume that there exist two equilibria, σ and τ, and that some players believe that σ will be played out, while others believe that τ will. Then, rather than an equilibrium, a mixture of σ and τ will be played out, so that $\mathbb{P}$ may not hold. Of course, this problem worsens as the number of players and/or equilibria grows.

The following mechanism, perhaps the first thing that comes to mind for our context, shows how big the problem of equilibrium selection can be for our goal.

> HOPE-FOR-THE-BEST: *Each player reports the valuations of all players (including himself). If all reports are the same, then (1) choose the state ω maximizing the sum of the reported valuations and (2) for each player i, choose the price P_i to be his reported value for ω (possibly minus a small discount ϵ to encourage i's participation). If not all reports coincide, then choose the "null outcome" (which all players are assumed to value 0) and price 0 for every player.*

It is trivial to see that the strategy profile in which each player reports all true valuations is a Nash equilibrium for HOPE-FOR-THE-BEST, indeed, it is the *truthful equilibrium.* It is also trivial to see that in this equilibrium the revenue is the maximum possible (disregarding the negligible quantity $n\epsilon$). Notice too, however, that HOPE-FOR-THE-BEST also has *additional* equilibria, $E_2, E_3, \ldots$, where in E_x all players report all true valuations divided by x. Thus, the truthful equilibrium is E_1, and in each E_x the utility of each player is increased by a factor x, and the money collected is a fraction $1/x$ of the maximum possible revenue. Accordingly,

- In the truthful equilibrium E_1 the designer is "happy", but the players are "sad", while
- in all other equilibria E_x the players are "happy" and the designer is "sad."

This being the case: *which equilibrium E_x is more likely to be selected?* Further, while each E_x at least maximizes social welfare, in plenty of other equilibria both revenue and social welfare are quite poor.[1] Given the multitude of available equilibria and the fact that different equilibria are preferable to different players: Will a play of HOPE-FOR-THE-BEST be an equilibrium and generate any revenue at all? In sum,

HOPE-FOR-THE-BEST *is extremely vulnerable to equilibrium selection.*

THE JPS MECHANISM . Notably, Jackson, Palfrey, and Srivastava [12] provided a quite different mechanism, but still yielding optimal revenue only at the truthful equilibrium τ. This time, however, τ is a much more meaningful equilibrium: *it is the only Nash equilibrium composed of weakly undominated strategies.* Somewhat counterintuitively, however, their solution too is vulnerable to equilibrium-selection. The point is that, as in HOPE-FOR-THE-BEST, there are plenty of equilibria σ that generate smaller revenue while being more attractive to *all* players. Again too, each such σ consists of reporting all true valuations divided by the same factor x. To be sure, this time each component σ_i is weakly dominated by some other strategy σ'_i. This means that, in all cases (i.e., for all possible subprofiles of strategies for the other players) σ_i provides no more utility to i than σ'_i does, while in at least some cases σ_i provides less utility to i than σ'_i. But in the JPS mechanism this happens in only one case: when *all other players "suicide"* (i.e., when all other players deliberately choose the worst possible strategies for themselves). Thus, as long as a single player does not believe that all others will commit mass suicide, all players prefer σ to the truthful and revenue-maximizing equilibrium τ. Accordingly,

the JPS mechanism too is very vulnerable to equilibrium selection.

[1] Let ω be any state such that $v_i(\omega) >> c > 0$ for all players i. And let σ be the strategy profile, where each σ_j consists of reporting that all players have the following valuation v: $v(\omega) = c$ and $v(x) = 0$ for any state $x \neq \omega$. Then, it is easy to see that σ is an equilibrium. Moreover, the revenue of σ is cn, and the social welfare of σ is $\sum_i v_i(\omega)$.

What has happened? Although “no one should want to play a weakly dominated strategy,” the problem is that the process of eliminating all *weakly* dominated strategies for yourself and the other players is not well defined. Unlike the iterated elimination of *strictly* dominated strategies, the iterated elimination of weakly dominated strategies depends on the order of elimination. For example, if one eliminates first “suicidal strategies” (in fact, if one eliminates first “suicide” for just another one of the players), then all equilibria become equally reasonable, and the attractive ones from the players’ point of view are those generating less revenue.

Collusion and Privacy *Collusion* and *privacy* can also prevent mechanisms from achieving their goals.

The problem of collusion in mechanism design is well recognized. The problem occurs for obvious reasons. Any equilibrium, even a dominant-strategy one, only guarantees that no *single* player has incentive to deviate from his strategy. However, two or more players may have all the incentive in the world to *jointly* deviate from their equilibrium strategies. Accordingly, by “guaranteeing” a property $\mathbb{P}$ at equilibrium, a classical mechanism is typically vulnerable to collusion. In a second-price auction, although the mechanism is dominant-strategy, if the players with the highest two valuations for the item on sale collude, then the revenue generated drops from the second-highest to the third-highest valuation. As for a more extreme example, Ausubel and Milgrom [1] show that two sufficiently informed players can totally destroy the economic efficiency of the famous VCG mechanism [22, 7, 9], although it too is dominant-strategy.

Privacy has been traditionally neglected in mechanism design, and considered a *quite separate* desideratum: nice to have perhaps, but not central for an incentive analysis. Yet, as especially argued by [10], it has a great potential to distort incentives, and thus to derail classical mechanisms from achieving their desired properties. A mechanism typically neglect privacy by requiring the players to reveal a lot of information about themselves. But if the players value privacy (which by definition implies that divulging their secret information causes them to receive a *negative utility*), then the mechanism gives them both positive and negative incentives, and it is no longer clear how these opposing forces will balance out.

Note that the JPS mechanism, HOPE-FOR-THE-BEST, and all traditional mechanisms *totally* disregard privacy by requiring the players to reveal all their information. Further, the JPS mechanism (unlike HOPE-FOR-THE-BEST) makes *no attempt to protect against collusion.* Indeed, it enables some pairs of players (i, j) to jointly deviate from the truthful equilibrium so as to improve the utility of i without hurting that of j. And when they so deviate *revenue cannot be maximum.*

1.4 Our Results

An Impossibility Result for Implementation in Dominant Strategies The problem of equilibrium selection fully disappears when a mechanism achieves its desired property $\mathbb{P}$ at a strictly dominant-strategy equilibrium, while still “lurks around” for weakly dominant-strategy equilibria. Unfortunately, we prove that neither strong nor weakly dominant-strategy mechanisms exist that can guarantee perfect revenue from perfectly informed players. Worse, our impossibility result holds even even if the mechanism designer were content to generate an *arbitrarily small fraction* of the optimal revenue. In sum, we prove the following.

Thm 1: *No weakly dominant-strategy mechanism guarantees a fraction ϵ of the optimal revenue.*

Thus, if we really want to achieve our goal without any equilibrium-selection (as well as privacy and collusion) problems, *the time has come to explore a different approach to mechanism design.*

A Possibility Result for Implementation in Robust Strategies Our main result is that our goal *can* be achieved, but by less classical means. We use mechanisms of extensive-form (that, as the Ascending English Auction, are played in several rounds), but adopt a new solution concept, *implementation in robust strategies* (*robust implementation* for short), as introduced by [5]. In our setting, where each player is perfectly informed about the others, their notion is equivalent to the following one.

Definition 1. *Let M be an extensive-form perfect-information mechanism, where the players may act simultaneously at some decision nodes. Process all its decision nodes in the following bottom-up fashion:*

- *At each decision node N of height 1, each player iteratively eliminates all strictly dominated strategies (for himself and all other players) for the normal-form subgame consisting of node N*
- *At every decision node N of height h, assuming recursively that all decision nodes of height $h-1$ have been already processed, the players iteratively eliminate all strictly dominated strategies for the subgame rooted at node N.*

We say that M robustly implements a property $\mathbb{P}$, if $\mathbb{P}$ holds for all profiles of strategies that survive this elimination procedure.

Notice that, although there are multiple ways to carry out the above elimination process, no rational player will ever play a strategy discarded by the above process. Thus M indeed guarantees $\mathbb{P}$ in a very *robust* sense, without in particular relying on any beliefs about the way the game could be played. Our second result can be summarized as follows

Thm 2: *There exists a robust-strategy mechanism $\mathcal{M}$ guaranteeing a fraction $1-\epsilon$ of the optimal revenue.*

Taken together, our impossibility and possibility results in particular yield a strong separation between implementation in robust strategies and implementation in dominant strategies.

Additional properties Our mechanism $\mathcal{M}$ actually satisfies even stronger properties: namely

- *The game yielded by our mechanism $\mathcal{M}$ has a* single *subgame-perfect equilibrium.*[2]
  We stress the word "single" because subgame-perfect equilibria, although more reasonable than Nash ones in extensive-form games, are not otherwise immune to equilibrium-selection problems.
  Perhaps interestingly, in the presence of collusion, our mechanism $\mathcal{M}$ has multiple ways to be truthful, but *only one* of them is a subgame-perfect equilibrium.
- *While guaranteeing perfect revenue, $\mathcal{M}$ also guarantees* perfect collusion resilience *and* perfect privacy.
  By saying that our mechanism is perfectly resilient to collusion we mean that $\mathcal{M}$ guarantees perfect revenue as long as not all players belong to the same coalition, and each coalition acts rationally. In our setting, a rational coalition maximizes the sum of the individual utilities of its members. (Only when the players have imperfect knowledge about each other, one may want to consider weaker models of coalition rationality.)
  By saying that $\mathcal{M}$ is perfectly private we essentially mean that in any rational play nothing can be learned about the players' valuations, except what is deducible from an outcome with perfect revenue. Of course, our $\mathcal{M}$ can be so "perfect" only because we are dealing with perfectly informed players (so that the only privacy concern is with respect to the designer/seller/auctioneer/outside world). But this *is* our setting, and thus one has both the right to demand and the obligation to deliver as a perfect solution as possible.

We stress that both properties above hold no matter how well collusive players cooperate. (In particular, they are free to make side-payments to each other and/or to enter into binding contracts with each other.)

Comparison with other work

- Note that our notion of collusion resiliency is stronger than that offered by other mechanisms. In particular, *group —or coalition— strategyproofness* [2, 16, 13, 18, 21] rules out collusion, but only under the assumption that the players are not able to make side payments to each other. Without restricting how players might cooperate, *t-truthful mechanisms* [8] offer protection against coalitions of at most t players, but only for single-value games. (In such games, a player i values some outcomes 0, and all other outcomes a fixed value v_i.) Again without restricting cooperation abilities, *collusion neutralization* [17, 5]

[2]For the non experts, this means that for each node and each player, only a single strategy survives the elimination process of Definition 1. That is, $\mathcal{M}$ is such that, at each node, every acting player has a single best action available to him.

offers collusion protection in more general games, but their notion too is weaker than the one considered in our paper. (Protection against the coalition of all players has also been considered and achieved, but only in Bayesian settings, where the distributions of player preferences are known to everyone, including the mechanism designer [14, 15, 3, 4].) Finally, a different approach altogether, *collusion leveraging,* has been submitted to this same conference [6].

- Some privacy preserving mechanisms have already appeared [20, 19, 10]. Their privacy, however, is either limited or gained by adding an additional layer to the mechanism —such as one or more mediators, envelopes, or encryption. By contrast, our mechanism $\mathcal{M}$ achieves perfect privacy without relying on any additional infrastructure. Indeed, $\mathcal{M}$ works by asking the players to take only *public actions.*

In Sum *To be really meaningful, mechanism design must seek mechanisms that are really robust. But robustly achieving even classical desiderata may require developing non classical approaches.*

2 Preliminaries

Contexts. In our paper we work with reasonably general contexts with semi-linear utilities. Namely, our context is defined by the following items:

- N, the finite set of *players:* $N = \{1, \ldots, n\}$
- $\Omega \times \mathbb{R}^n$, the set of possible outcomes, where Ω is finite.
  A member ω of Ω is referred to as a *state* and a member P of $\mathbb{R}^n$ is referred to as a *price profile.* Set Ω is required to include the *empty state*, denoted by $\perp$.
- V is the set of all possible profiles of (non-negative) player *types* or *valuations.*
  Each type is a function, from the set of states to the set $\mathbb{N}$ of natural numbers, mapping $\perp$ to 0.
  We consistently denote by TV the profile of the *true types* (that is, for each player i, TV_i describes i's actual value for each possible state).
- u_i, for each player i, is i's *utility function*, mapping outcomes to real numbers as follows: $u_i(\omega, P) = TV_i(\omega) - P_i$. That is, i's utility is i's true value for the state minus the price he pays.
- $\mathbb{C}$, a partition of N specifying who colludes with whom.
  If S is a subset in $\mathbb{C}$, then S is the maximal subset of players colluding with each other. A player i is independent if $\{i\}$ is in $\mathbb{C}$, and the context is non-collusive if all players are independent.

Accordingly, to specify a context C, it suffices to specify just its "variable" components: that is, the quadruple $(N, \Omega, TV, \mathbb{C})$. If the context is non-collusive, it suffices to specify the triple (N, Ω, TV). Each player i knows his own type. Each independent player tries to maximize his own utility function. Each collusive set, that is a subset in $\mathbb{C}$ with cardinality greater than 1, tries to maximize the sum of the utilities of its members.

We say that a context is *perfect-knowledge*, equivalently that the players are perfectly informed, if the entire true-valuation profile TV (as well as the partition $\mathbb{C}$ if the context is collusive) is common knowledge to all players. We stress that the mechanism *designer* has no knowledge about TV (or $\mathbb{C}$)! In other words, we adhere to the classic spirit of mechanism design, where all knowledge lies with just the players.

Strategies and Mechanisms. We now must specify the players'strategies, and how these lead to outcomes. Traditionally, attention is restricted to mechanisms in which each player, simultaneously with the others, announces a type for himself (which may or may not coincide with his true valuation function). For such mechanisms, thanks to the *revelation principle*, a player's set of strategies consists of the set of all possible valuations.

In our case, however, the players do not only know their own types, but also those of the others. And to leverage this extra knowledge, it is crucial that the players be able to announce types for all players. That is, a player's *strategy* consists of a profile of valuations (in other words, it is a member of V). The *empty strategy* is the one whose valuations map every possible state to 0.

A *mechanism* for a context (N, Ω, TV) consists of a (possibly probabilistic) function $M : V^n \to \Omega \times \mathbb{R}^n$ satisfying the following

> **Opt-Out Condition:** For any strategy profile $v = v_1, \ldots, v_n$, if $M(v) = (\omega, P)$ then $P_i = 0$ whenever v_i is the empty strategy.

PLAYS. A *play* σ of a mechanism M consists of a profile of strategies. If M is probabilistic, then $M(\sigma)$ is a distribution over outcomes, and $u_i(M(\sigma))$ is the expected utility of player i over such distribution, that is, it is short hand for $\mathbb{E}[u_i(M(\sigma))]$.

SOCIAL WELFARE, REVENUE, AND OUR GOAL. The *social welfare* and the *revenue* of an outcome (ω, p) are respectively defined to be $\sum_i TV_i(\omega)$ and $\sum_i p_i$.

The *maximum rational revenue* for a context $C = (N, \Omega, TV)$ is defined to coincide with the *maximum social welfare* (MSW for short), that is, $\max_\omega \sum_i TV_i(\omega)$.

We are interested in designing mechanisms (essentially) guaranteeing the maximum rational revenue.

3 Impossibility Result for DST mechanisms

Let us prove that DST mechanisms are incapable of properly leveraging external knowledge: namely, in a perfect-knowledge context, they cannot guarantee even a minuscule fraction of the maximum rational revenue.

Definition 2. *A DST mechanism M guarantees a fraction ϵ of the maximum rational revenue if for any context $C = (N, \Omega, TV)$ we have*

$(*)$ $\quad M(TV, \ldots, TV) = (x, P)$ *implies* $\sum P_i \geq \epsilon \cdot MSW$.

Note that, in proposition $(*)$, each TV is not just the true valuation of a single player, but the profile of all such valuations, because a player's strategy includes his declaration about the others' valuations as well.

Note too that the mechanism is not required to choose the outcome which maximizes the social welfare. Moreover, when not all the players are telling the truth, there is no requirement on the behavior of the mechanism.

Finally note the following immediate corollary of the opt-out condition. Namely,

> *Non-negative utility property:* if M is a DST mechanism and $M(v^1, \ldots, v^n) = (\omega, P)$, then $P_i \leq v_i^i(\omega)$.

Theorem 1. *For any $\epsilon > 0$ no DST mechanism M guarantees a fraction ϵ of the maximum rational revenue.*

Proof. We actually prove our result even for contexts with just two players and only two possible outcomes. Without loss of generality, consider the context (N, Ω, TV) where $N = \{1, 2\}$ and $\Omega = \{\perp, \omega\}$. In such a context, a valuation v_i of a player i coincides with a single number $v_i(\omega)$ (because $v_i(\perp)$ is bound to be 0), and so a strategy v for i coincides with a pair of numbers, $v = (c_1, c_2)$, where c_1 is the declared value for player 1 and c_2 the declared value for player 2.

Our proof is by contradiction. We start by analyzing the behavior of M when the two players make identical and positive (but not necessarily truthful) declarations. More precisely, we prove the following proposition:

$(\star)$ $\quad$ if $c_1, c_2 > 0$, then $M(\ (c_1, c_2)\ ,\ (c_1, c_2)\ ) = (x, (P_1, P_2))$ where

$\star_1$: $\quad P_1 + P_2 \geq \epsilon \cdot (c_1 + c_2)$

$\star_2$: $\quad x = \omega$

To see that proposition $(\star)$ holds, assume the players bid truthfully; that is assume that $c_1 = TV_1(\omega)$ and $c_2 = TV_2(\omega)$. In this case, according to $(*)$ the mechanism must extract a revenue of at least $\epsilon \cdot MSW = \epsilon \cdot (c_1 + c_2)$, and thus $P_1 + P_2 \geq \epsilon \cdot (c_1 + c_2)$, in agreement with inequality $\star_1$.

Now, the hypothesis $c_1 + c_2 > 0$ implies $P_1 + P_2 > 0$. Thus, in light of the non-negative utility property, the state returned by M cannot be $\perp$. Since ω is the only other state, M has to return ω in agreement with equality $\star_2$.

Consider now the declaration $K = (1, 1)$ and let $M(K, K) = (y, Q)$. Then proposition $(\star)$ guarantees that $y = \omega$ and that $Q_1 + Q_2 \geq 2\epsilon$. This implies that $Q_i \geq \epsilon$ for at least a player i. Thus, without loss of generality, we can assume $Q_1 \geq \epsilon$.

Consider now the strategy $\tilde{K} = (\epsilon/2, \epsilon/2)$, and let us analyze the behavior of $M(\tilde{K}, K)$. Let $M(\tilde{K}, K) = (x, P)$.

We start by proving that $x = \omega$. Assume for contradiction purposes that $x = \perp$. Then, when $TV = K$ (and thus player 1 is not truthful), player 2 has an incentive to lie. Indeed, by being truthful, under the current assumption, his utility is 0. However, if player 2 chose the strategy $\tilde{K}$, then according to $(\star)$, the outcome would be (ω, P_1, P_2). In this case, according to the non-negative utility property, since player 2's self-valuation is $\epsilon/2$, $P_2 \leq \epsilon/2$. Thus player 2's utility would be at least $1 - \epsilon/2$. Since this utility is positive, while his utility of being truthful is 0, player 2 has an incentive to lie when $TV = K$ and player 1's strategy is $\tilde{K}$. Therefore we must have $x \neq \perp$, or equivalently $x = \omega$.

Let us now analyze the possible values for P_1 and derive a contradiction in every case.

1. *Case 1:* $P_1 < \epsilon$. In this case, assume that $TV = K$ and compute player 1's utility under the following two strategy profiles: (K, K) and $(\tilde{K}, K)$. In the first case we already know that $M(K, K) = (\omega, Q)$, where $Q_1 \geq \epsilon$. Therefore player 1's utility when being truthful is $1 - Q_1$ which is at most $1 - \epsilon$. On the other hand, under the strategy profile $(\tilde{K}, K)$, player 1's utility is equal to $1 - P_1$ and thus strictly greater than $1 - \epsilon$ by hypothesis. Thus, the context $(\{1, 2\}, \{\perp, \omega\}, K)$ contradicts the dominant-strategy truthfulness of M.

2. *Case 2:* $P_1 > \epsilon/2$. In this case, since $M(\tilde{K}, K) = (\omega, P)$ and $\tilde{K} = (\epsilon/2, \epsilon/2)$, the non-negative utility property implies that $P_1 \leq \epsilon/2$, and thus a contradiction.

In sum, if M guarantees an ϵ fraction of the maximum possible revenue, no price profile exists for $M(\tilde{K}, K)$ that does not contradict the dominant-strategy truthfulness of M. Since we have not assumed any property of M beyond its being DST, this establishes our theorem. *Q.E.D.*

4 Our Mechanism

Notation In the mechanism below,

- ϵ and ϵ^i_j, for $i \in \{2, \ldots, n\}$ and $j \in \{1, \ldots, n\}$, are constants such that
$$\frac{1}{5n} > \epsilon > \epsilon^2_1 > \cdots > \epsilon^2_n > \epsilon^3_1 > \ldots > \epsilon^3_n > \cdots > \epsilon^n_1 > \cdots > \epsilon^n_n > 0.$$
- Numbered steps are taken by the players, while steps marked by letters are taken by the mechanism.
- Sentences between quotation marks are comments, and could be excised if no clarification is needed.
- We denote by n_r the number of outcomes (ω, P) with revenue r. For all such outcomes, we denote by $0 \leq f_r(\omega, P) < n_r$ the rank of the outcome (ω, P) in the lexicographic order that first considers the state and then the price profile (where $P_1, \ldots P_n$ precedes $P'_1 \ldots P'_n$ whenever $P_1 > P'_1$, etc.).

Mechanism $\mathcal{M}$

(1) *Player* 1 *announces a state* $\omega^\star$ *and a profile* K^1 *of natural numbers.*

"$(\omega^\star, K^1)$ is player 1's proposed outcome, allegedly an outcome of maximum revenue."

(a) *Set $\omega = \perp$, and $P_i = 0\ \forall i$. If $\sum_i K_i^1 = 0$, the mechanism ends right now. Otherwise, proceed to Step 2.*

"Whenever the mechanism ends, ω and P will be, respectively, the final state and price profile."

(2,…, n) *In Step i, $2 \leq i \leq n$, player i publicly announces a profile Δ^i of natural numbers such that $\Delta_i^i = 0$.*

"By so doing i suggests to raise the current price of j, that is $K_j^1 + \sum_{\ell=2}^{i-1} \Delta_j^\ell$, by the amount Δ_j^i."

(b) *For each player i, publicly select bip_i and $P_i^\star$ as follows. Let $R_i = \{j\ :\ \Delta_i^j > 0\}$.*
If $R_i \neq \emptyset$, then bip_i is highest player in R_i, and $P_i^\star = K_i^1 + \sum_{\ell=2}^{bip_i} \Delta_i^\ell$. Else, $bip_i = 1$ and $P_i^\star = K_i^1$.

"We refer to bip_i as the best informed player about i, and to $P_i^\star$ as the provisional price of i."

($n+1$) *Each player i such that $P_i^\star > 0$ simultaneously announces YES or NO.*
By default, each player i such that $P_i^\star = 0$ announces YES, and player 1 announces YES if $bip_1 = 1$.

"Each player i announces YES or NO to $\omega^\star$ as the final state and to $P_i^\star - \epsilon$ as his own price. (By default player 1 accepts his own price if no one raises it.) At this point the players are done playing, and the mechanism proceeds as follows. If all say YES, the updated proposal $(\omega^\star, P^\star)$ is implemented with probability 1. Else:

- With very high probability the null outcome is chosen, except that the best-informed players of those saying NO are punished.
- With small probability the null outcome is chosen
- With very small probability, proportional to the number of players saying YES, we implement $(\omega^\star, P^\star)$ as if all said YES.

Importantly, as we shall see, all get a small reward at the end for their knowledge."

(c) *Let Y be the number of players announcing YES. If $Y = n$, then reset ω to $\omega^\star$ and each P_i to $P_i^\star - \epsilon$, and go to Step g. If $Y < n$, proceed to Step d.*

(d) *Publicly flip a biased coin c_1 such that $\Pr[c_1 = Heads] = 1 - \epsilon$.*

(e) *If $c_1 = Heads$, reset P_{bip_i} to $P_{bip_i} + 2P_i^\star$ for each player i announcing NO.*

(f) *If $c_1 = Tails$, letting $B = \sum_{i \text{ announces NO}} P_i^\star$, flip a biased coin c_2 such that $\Pr[c_2 = Heads] = \frac{Y}{nB}$.*
If $c_2 = Heads$, reset ω to $\omega^\star$ and each P_i to $P_i^\star - \epsilon$.
If $c_2 = Tails$, ω and P continue to be $\perp$ and 0^n.

(g) *Reset P_1 to $P_1 - \epsilon - 2\epsilon \sum_j K_j^1 + \epsilon \frac{f_r(\omega)}{n_r}$ and each other P_i to $P_i - \epsilon - \sum_j \epsilon_j^i \Delta_j^i$.*

"Although players' prices may be negative, we prove that the mechanism never loses money, and that in the unique rational play the utility of every player is non-negative. For clarity, our rewards are proportional to prices and raises."

5 Analysis of Our Mechanism

Mechanism $\mathcal{M}$ induces a game G whose game tree has height $n+1$, and where only players act at each internal node. (The mechanism tosses all its coins at leaf nodes, that are defined to be of height 0.) At each node of height 1 all players act simultaneously, and at every other internal node only a single player acts. Specifically, at each node of height $h \geq 2$ the only acting player is player

$$i_h \triangleq n - h + 2.$$

For each internal node N, we denote by G^N the subgame rooted at N. Recall that a strategy σ_i of player i in G specifies, for each node at which i acts, which action i chooses among all those available to him. By σ_i^N we denote the restriction of σ_i to subgame G^N. Given a restricted strategy profile σ^N for G^N, the outcome of $\mathcal{M}$ obtained by executing σ^N is denoted by $\mathcal{M}(\sigma^N)$.

For uniformity, we find it sometimes convenient to assume that every player i belongs to a (necessarily unique) collusive set, denoted by $\mathscr{C}_i$. If i is independent, then $\mathscr{C}_i = \{i\}$.

5.1 Statements of Our Lemmas

Lemma 1. *If N is a node of height 1, then G^N has a unique subgame-perfect equilibrium σ^N, where*

- *If i is independent, then σ_i^N consists of announcing YES if and only if $TV_i(\omega^\star) \geq P_i^\star$.*
- *If i belongs to a coalition $\mathscr{C}$, then σ_i^N consists of announcing YES if and only if*

$$bip_i \in \mathscr{C} \quad \text{or} \quad \sum_{j\in\mathscr{C}} TV_j(\omega^\star) \geq \sum_{j\in\mathscr{C}} P_j^\star.$$

The proof of this lemma is based on the fact that the probability that an outcome is executed is monotone with the number of players who announce YES. Thus, it is strictly dominant to announce YES, if and only if the player has positive utility from this outcome and price.

Lemma 2. *Let N be a node of height $h \in [2, n]$, $i = i_h$, and $\mathscr{C} = \mathscr{C}_i$. If every player x plays his unique subgame-perfect strategy σ_x^M at each proper subgame G^M of G^N, then G^N has a unique subgame-perfect equilibrium where i acts as follows at node N: For each collusive set $\mathscr{D} \neq \mathscr{C}$,*

1. *if*
$$\sum_{j\in\mathscr{D}} \left(K_j^1 + \sum_{\ell=2}^{i-1} \Delta_j^\ell \right) \geq \sum_{j\in\mathscr{D}} TV_j(\omega^\star)$$
then i announces $\Delta_j^i = 0$ for all $j \in \mathscr{D}$;
2. *if*
$$\sum_{j\in\mathscr{D}} \left(K_j^1 + \sum_{\ell=2}^{i-1} \Delta_j^\ell \right) < \sum_{j\in\mathscr{D}} TV_j(\omega^\star)$$
then letting k be the minimal player in $\mathscr{D}$, i announces $\Delta_j^i = 0$ for all $j \in \mathscr{D} \setminus \{k\}$ and
$$\Delta_k^i = \sum_{j\in\mathscr{D}} \left(TV_j(\omega^\star) - K_j^1 - \sum_{\ell=2}^{i-1} \Delta_j^\ell \right).$$

For his own collusive set $\mathscr{C}$,

1. *if*
$$\sum_{j\in\mathscr{C}} \left(K_j^1 + \sum_{\ell=2}^{i-1} \Delta_j^\ell \right) \geq \sum_{j\in\mathscr{C}} TV_j(\omega^\star) \quad \text{or} \quad \text{it is the case that } k \in \mathscr{C} \text{ for all } k > i,$$
then i announces $\Delta_j^i = 0$ for all $j \in \mathscr{C}$;
2. *if*
$$\sum_{j\in\mathscr{C}} \left(K_j^1 + \sum_{\ell=2}^{i-1} \Delta_j^\ell \right) < \sum_{j\in\mathscr{C}} TV_j(\omega^\star) \quad \text{and} \quad \text{there exists player } j > i \text{ such that } j \notin \mathscr{C},$$
then letting k be the minimal player in $\mathscr{C} \setminus \{i\}$, i announces
$$\Delta_k^i = \sum_{j\in\mathscr{C}} \left(TV_j(\omega^\star) - K_j^1 - \sum_{\ell=2}^{i-1} \Delta_j^\ell \right).$$

This lemma is technically involved, but conceptually simple. First, we show that a player i never wants to "overbid," that is raise the price of another player j to more than j's true valuation for the proposed state $\omega^\star$. When j is independent, this holds because we know that j will announce NO to any price above his true valuation, and thus no player after i will want to further raise j's price. Therefore, overbidding on j will cause i to be punished. Care must still be taken to verify the Step-g rewards of i and j will not change this simple analysis. (For example j will not accept a higher price in order to get more reward for volunteering his knowledge about other players.) For coalitions, the argument is more subtle.

After ruling out overbidding, we also show that a player i never wants to "underbid," that is not raise the price of a player j when it is below j's true valuation for the proposed state. Again, this is easier to argue for independent players. Arguing this point for coalitions is the only time that requires exploiting the n^2 reward values $\epsilon_{i,j}$.

Lemma 3. *Let N be the root of the tree (so that $G^N = G$) and let every player x play his unique subgame-perfect strategy σ_x^M at each proper subgame G^M of G. Then G has a unique subgame-perfect equilibrium where player 1 acts as follows at node N:*

1. *player 1 announces $\omega^\star$, the lexicographically first state ω such that $\sum_\ell TV_\ell(\omega) = MSW$;*
2. *for each collusive set $\mathscr{D}$, letting i be the minimal player in $\mathscr{D}$, player 1 announces $K_i^1 = \sum_{j\in\mathscr{D}} TV_j(\omega^\star)$, and $K_j^1 = 0$ for each $j \in \mathscr{D} \setminus \{i\}$.*

The proof of this lemma is also done in two stages. First, given Lemma 2, we prove that it is dominant for player 1 to set the prices correctly (although not exactly truthfully in the case of a coalition). Finally, as the prices are set correctly, choosing the outcome which maximizes the total welfare dominates any other course of action.

Proofs of the three lemmas will come in the final version.

5.2 Our Main Theorem

Theorem 2. *Let σ be the unique subgame perfect equilibrium of G, and let $(\omega, P) = \mathcal{M}(\sigma)$. Then:*

(1) $\sum_i TV_i(\omega) = MSW$, and

(2) $\sum_i P_i \geq (1 - 4\epsilon n)MSW$.

Proof. In execution σ, by Lemma 3, player 1 announces $\omega^\star$ such that $\sum_\ell TV_\ell(\omega^\star) = MSW$ and, for each coalition $\mathscr{D}$, also announces $K_\ell^1 = \sum_{j\in\mathscr{D}} TV_j(\omega^\star)$, where ℓ is the minimal player in $\mathscr{D}$. Thus $\sum_i K_i^1 = MSW$.

If $MSW = 0$, then $\sum_i K_i^1 = 0$ and $\mathcal{M}$ ends at Step a, with $\omega = \bot$ and $P_i = 0$ for each player i. Therefore $\sum_i TV_i(\omega) = \sum_i TV_i(\bot) = 0 = MSW$ and $\sum_i P_i = 0 = MSW$.

If $MSW > 0$, then $\sum_i K_i^1 > 0$ and $\mathcal{M}$ ends at Step g. By Lemma 2, for each player $i \neq 1$, i announces $\Delta_k^i = 0$ for each k. Therefore for each player i, $bip_i = 1$. Furthermore, the total price for each coalition $\mathscr{D}$ equals $\mathscr{D}$'s total true valuation for $\omega^\star$: that is, $\sum_{\ell\in\mathscr{D}} P_\ell^\star = \sum_{\ell\in\mathscr{D}} K_\ell^1 = \sum_{\ell\in\mathscr{D}} TV_\ell(\omega^\star)$. By Lemma 1, every player in $\mathscr{D}$ announces YES in Step $n+1$. This implies that, at the end of Step c we have $Y = n$, $\omega = \omega^\star$, and, for each coalition $\mathscr{D}$, $\sum_{\ell\in\mathscr{D}} P_\ell = \sum_{\ell\in\mathscr{D}} P_\ell^\star - |\mathscr{D}|\epsilon = \sum_{\ell\in\mathscr{D}} TV_\ell(\omega^\star) - |\mathscr{D}|\epsilon$. Because $Y = n$, the execution of $\mathcal{M}$ will then proceed directly to Step g, which does not reset the current state. Thus we have that

$$\sum_i TV_i(\omega) = \sum_i TV_i(\omega^\star) = MSW.$$

Because the reward given to each player $i > 1$ in Step g is ϵ, and player 1 gets at most $\epsilon + 2\epsilon MSW$, then the final revenue of the mechanism is

$$\sum_i P_i > \left(\sum_i TV_i(\omega^\star) - n\epsilon\right) - (n-1)\epsilon - \epsilon - 2\epsilon MSW > (1 - 4\epsilon n)MSW,$$

where we parenthesized the prices after step c, and used that MSW is integer and thus $MSW \geq 1$. *Q.E.D.*

References

[1] L.M. Ausubel and P. Milgrom. The Lovely but Lonely Vickrey Auction. *Combinatorial Auctions*, MIT Press, pages 17-40, 2006.

[2] S. Barbera and M.O. Jackson. Strategy-Proof Exchange. *Econometrica*, Vol.63, No.1, pages 51-87, Jan., 1995.

[3] Y. Che and J. Kim. Robustly Collusion-Proof Implementation. *Econometrica*, Vol.74, No.4, pages 1063-1107, Jul., 2006.

[4] Y. Che and J. Kim. Optimal collusion-proof auctions. *Journal of Economic Theory*, Vol.144, Issue 2, pages 565-603, Mar. 2009.

[5] J. Chen and S. Micali. A New Approach to Auctions and Resilient Mechanism Design. *Symposium on Theory of Computing*, pages 503-512, 2009. Full version available at http://people.csail.mit.edu/silvio/Selected Scientific Papers/Mechanism Design/.

[6] J. Chen S. Micali, and P. Valiant. Levaraging Collusion in Combinatorial Auctions. Submitted to ICS2010.

[7] E.H. Clarke. Multipart Pricing of Public Goods. *Public Choice*, Vol.11, No.1, pages 17-33, Sep., 1971.

[8] A. Goldberg and J. Hartline. Collusion-Resistant Mechanisms for Single-Parameter Agents. *Symposium on Discrete Algorithms*, pages 620-629, 2005.

[9] T. Groves. Incentives in Teams. *Econometrica*, Vol. 41, No. 4, pages 617-631, Jul., 1973.

[10] S. Izmalkov, M. Lepinski, and S. Micali. Perfect Implementation of Normal-Form Mechanisms. *MIT-CSAIL-TR-2008-028*, May 2008. (A preliminary version appeared in FOCS'05, pages 585-595, with title "Rational Secure Computation and Ideal Mechanism Design".)

[11] M. Jackson. Implementation in Undominated Strategies: a Look at Bounded Mechanisms. *Review of Economic Studies*, Vol. 59, No. 201, pages 757-775, Oct., 1992.

[12] M.O. Jackson, T. Palfrey and S. Srivastava. Undominated Nash Implementation in Bounded Mechanisms. *Games and Economic Behavior*, 1994.

[13] K. Jain and V. Vazirani. Applications of Approximation Algorithms to Cooperative Games. *Symposium on Theory of Computing*, pages 364-372, 2001.

[14] J.J. Laffont and D. Martimort. Collusion under Asymmetric Information. *Econometrica*, Vol.65, No.4, pages 875-911, Jul. 1997.

[15] J.J. Laffont and D. Martimort. Mechanism Design with Collusion and Correlation. *Econometrica*, Vol. 68, Issue 2, pages 309-342, Mar. 2000.

[16] H. Moulin. Incremental cost sharing: Characterization by coalition strategy-proofness. *Social Choice Welfare*, Vol.16, No.2, pages 279-320, Feb. 1999.

[17] S. Micali and P. Valiant. Resilient Mechanisms for Unrestricted Combinatorial Auctions, submitted to FOCS 2007; SODA 2008; and STOC 2009. Available at http://people.csail.mit.edu/silvio/Selected Scientific Papers/Mechanism Design

[18] H. Moulin and S. Shenker. Strategyproof Sharing of Submodular Costs: Budget Balance Versus Efficiency. *Economic Theory*, Vol.18, No.3, pages 511-533, 2001.

[19] F. McSherry and K. Talwar. Mechanism Design via Differential Privacy.*FOCS07*, 2007.

[20] M. Naor, B. Pinkas and R. Sumner. Privacy Preserving Auctions and Mechanism Design, *EC99*, 1999.

[21] M. Pal and E. Tardo, Group Strategyproof Mechanisms via Primal-Dual Algorithms. *44th Annual IEEE Symposium on Foundations of Computer Science (FOCS'03)*, pages 584-593, 2003.

[22] W. Vickrey. Counterspeculation, Auctions, and Competitive Sealed Tenders. *The Journal of Finance*, Vol. 16, No. 1, pages 8-37, Mar., 1961.