A Preliminary Introduction to Mechanism Design Theory

Hongbin Cai and Xi Weng

Department of Applied Economics, Guanghua School of Management

Peking University

October 2014

Contents

1 Introduction

1. The design of the institutions matters! What are institutions? *Rules of games!*

   Institutions are the rules of the game in a society or, more formally, are the humanly devised constraints that shape human interaction. . . . In the jargon of the economist, institutions define and limit the set of choices of individuals. (North 1990)

   The 2000-2001 "3G" mobile-phone license auctions not only raised one hundred billion dollars and attracted intense media scrutiny, they also provide an excellent illustration of our points about practical design. Even though the licenses sold were very similar in each of the nine west European auctions, the different auction designs resulted in revenues that varied from less than 20 dollars per capita in Switzerland to almost 600 dollars per capita in the United Kingdom. (Klemperer 2004)

2. What is mechanism theory? *Design of institutions that satisfy certain objectives, assuming that the individuals interacting through the institution will act strategically and may hold private information that is relevant to the decision at hand.*

3. In a mechanism, each individual has a message (or strategy) space and decisions result as a function of the messages chosen.

2 A General Mechanism Design Setting

1. Individuals: A finite group of individuals interact. This set is denoted $N = \{1, 2, \cdots, n\}$.

2. Decisions: The set of potential social decisions is denoted D. The set of decisions may

be finite or infinite depending on the application.

3. Preferences and Information: Each individual i holds private information, which is represented by a type $\theta_i \in \Theta_i$. Let $\theta = (\theta_1, \cdots, \theta_n) \in \Theta = \Theta_1 \times \cdots \Theta_n$. Individuals have preferences over decisions that are represented by a utility function $v_i : D \times \Theta_i \to \mathbb{R}$. So, $v_i(d; \theta_i)$ denotes the benefit that individual i of type $\theta_i \in \Theta_i$ receives from a decision $d \in D$.

4. Social Choice Function: a mapping $f : \Theta \to D$. A decision rule $f(\cdot)$ is efficient if

$$\sum_i v_i(f(\theta), \theta_i) \geq \sum_i v_i(d', \theta_i)$$

for all θ and d'.

Example 1 *Allocating a Private Good. An indivisible good is to be allocated to one member of society. For instance, the rights to an exclusive license are to be allocated or an enterprise is to be privatized. Each individual's valuation of the good is θ_i. The decision set is $D = \{d \in \{0, 1\}^n : \sum_i d_i = 1\}$ where $d_i = 1$ means that individual i gets the good. The preference is $v_i(d, \theta_i) = d_i \theta_i$.*

Example 2 *A Public Project. A society with n agents is deciding on whether or not to build a public project at a cost c. The cost of the public project is to be equally divided. Here $D = \{0; 1\}$ with 0 representing not building the project and 1 representing building the project. Individual i's value from use of the public project is represented by θ_i. The utility* *function of i can then be represented as*

$$v_i(d, \theta_i) = d(\theta_i - c/n).$$

Questions: What social choice functions can be implemented when each individual's preference is private information? Use institutions/mechanisms to implement a social choice function. A mechanism is a pair $\{M, g\}$, where $M = M_1 \times \cdots \times M_n$ is a cross product of message or strategy spaces and $g : M \to D$ is an outcome function. Thus, for each profile of messages $m = (m_1, \cdots, m_n)$, $g(m) \in D$ represents the resulting decision. Once the preferences of the individuals are specified, then a mechanism induces a Bayesian game with incomplete information where each individual's action set is M_i. Note that a mechanism is NOT a game but specifies the rules of a game!

Given a social choice function $f(\cdot)$, we say that mechanism $\{M, g\}$ implements f if there exists an equilibrium strategy profile $(s_1^\star, \cdots, s_n^\star)$ of the Bayesian game induced by $\{M, g\}$ such that

$$g(s_1^\star(\theta_1), \cdots, s_n^\star(\theta_n)) = f(\theta_1, \cdots, \theta_n)$$

for all $(\theta_1, \cdots, \theta_n) \in \Theta_1 \times \cdots \times \Theta_n$. The mechanism design problem is to design a mechanism so that when individuals interact through the mechanism, they have incentives to choose messages as a function of their private information that leads to socially desired outcomes.

3 Dominant Strategy Mechanism Design

3.1 Dominant Strategies

A strategy profile $(s_1^\star, \cdots, s_n^\star)$ constitutes a (weakly) dominant strategy equilibrium of mechanism $\{M, g\}$ if for all i and θ_i,

$$v_i(g(s_i^\star(\theta_i, m_{-i})), \theta_i) \geq v_i(g(m_i', m_{-i}), \theta_i).$$

A social choice function $f(\cdot)$ is implemented in dominant strategies by the mechanism$\{M, g\}$ if there exist strategies $s_i^\star : \Theta_i \to M_i$ such that $s_i^\star$ is a dominant strategy for each individual i.

A dominant strategy has the strong property that it is optimal for a player no matter what the other players do. When dominant strategies exist, they provide compelling predictions for strategies that players should employ. Also dominant strategy implementation does not require any knowledge about the joint distribution of private information $(\theta_1, \cdots, \theta_n)$.

3.2 Direct Mechanism and the Revelation Principle

The set of all possible mechanisms is very large. The revelation principle guarantees that we can restrict attention to very simple mechanisms called direct mechanisms.

Definition 1 *A direct mechanism is a mechanism in which $M_i = \Theta_i$ and $g(\theta) = f(\theta)$. The social choice function f is truthfully implementable in dominant strategies (or dominant strategy incentive compatible) if $s_i^\star(\theta_i) = \theta_i$ for all i and θ_i is a dominant strategy equilibrium of the direct mechanism.*

The Revelation Principle for Dominant Strategies: If a mechanism $\{M, g\}$ implements a social choice function f in dominant strategies, then f is truthfully implementable in dominant strategies.

3.3 The Gibbard-Satterthwaite Theorem

We say a social choice function is dictatorial if there exists an individual i such that f always picks i's top-ranked choice:

$$f(\theta) \in \text{argmax}_{d \in D} v_i(d, \theta_i).$$

Theorem 1 *Suppose that D is finite and type spaces include all possible strict orderings over D.*[1] *A social choice function with at least three elements in its range is truthfully implementable in dominant strategies if and only if it is dictatorial.*

The Gibbard-Satterthwaite theorem has quite negative implications for the hopes of implementing non-trivial decision rules in dominant strategies in a general set of environments. It implies that transfer functions will be needed for dominant strategy implementation of non-dictatorial decision rules in some settings. Before discussing the role of transfer functions, let us point out some prominent settings where the preferences do not satisfy the richness of types assumption of the Gibbard-Satterthwaite theorem and there exist non-dictatorial social choice functions truthfully implementable in dominant strategies.

[1]We say that type spaces include all possible strict orderings if for any ordering of D, there exist θ_i such that $v_i(d, \theta_i) > v_i(d', \theta_i)$ if d is ordered before d'.

3.4 Single-Peaked Preferences

Single peaked preference domains are widely used in modeling in voting games. In such a setting, the decision set has a single dimension: $D \subset \mathbb{R}$. Individuals have single-peaked preferences over$D$ if for each i and θ_i, there exists $p(\theta_i) \in D$, called the peak of i's preferences, such that $p(\theta_i) \geq d > d'$ or $p(\theta_i) \leq d < d'$ imply that $v_i(d, \theta_i) > v_i(d', \theta_i)$.

In a single peaked preference domain, there are social choice functions truthfully implementable in dominant strategies. For instance, consider the social choice function which requires each individual to declare their peak and then selecting the median (with a tie-break in the case of an even number of individuals). This results in truthful announcements of peaks as a dominant strategy. The same is true of variations on the median voting rule, such as taking the maximum of the peaks, the minimum, or any order statistic.

3.5 Transfers and Groves Mechanism

In order to provide the incentives necessary to implement the desirable social choice funtion, it may be necessary to tax or subsidize various individuals. The transfer functions are denoted by $t : \Theta \to \mathbb{R}^n$. The function $t_i(\theta)$ represents the payment that i receives (or makes if it is negative) based on the announcement of θ (notice that we are using direct mechanism now).

A transfer function t is said to be feasible if $0 \geq \sum_i t_i(\theta)$ for all θ and is balanced if $0 = \sum_i t_i(\theta)$.

Each individual's preference is quasi-linear in the transfers. The utility that i receives if

$\hat{\theta}$ is the "announced" vector of types and i's true type is θ_i is

$$u_i(\hat{\theta}, \theta_i, d, t) = v_i(d(\hat{\theta}), \theta_i) + t_i(\hat{\theta}).$$

With transfers, a mechanism can be redefined as a pair $\{M, g\}$, where $M = M_1 \times \cdots \times M_n$ is a cross product of message or strategy spaces and $g : M \rightarrow D \times \mathbb{R}^n$ represents the resulting decision and transfers.

It is shown by Theodore Groves that an efficient social choice function can be truthfully implementable in dominant strategies using transfers.

Theorem 2 *If f is an efficient social choice function and for each i there exists a function $x_i : \times_{j \neq i} \Theta_j \rightarrow \mathbb{R}$ such that*

$$t_i(\theta) = x_i(\theta_{-i}) + \sum_{j \neq i} v_j(f(\theta), \theta_j),$$

then f is truthfully implementable in dominant strategies using transfers t.

The core idea of the Groves mechanism is to make each individual's transfer function take into account the marginal social impact on other individuals made by his announcement of θ_i (internalize the externality). When looking at this social impact together with his own selfish utility, the individual has exactly the total social value in mind when deciding on a strategy. This leads to efficient decision making.

3.6 The Pivotal Mechanism and Vickrey Auctions

One simple version of the Groves mechanism is called pivotal mechanism proposed by Clarke. Let $x_i(\theta_{-i}) = -\max_{d\in D}\sum_{j\neq i} v_j(d,\theta_j)$. In this case, i's transfer becomes

$$t_i(\theta) = \sum_{j\neq i} v_j(f(\theta),\theta_j) - \max_{d\in D}\sum_{j\neq i} v_j(d,\theta_j) \leq 0.$$

The pivotal mechanism is always feasible. Individual i makes a payment if and only if his type is pivotal: i's presence makes a difference in the maximizing choice of d. The pivotal mechanism reduces to a well-known auction form in the context of the allocation of indivisible objects. In that context there is a simple auction that is dominant strategy incentive compatible, as first noticed by William Vickrey. It turns out that this auction form, commonly referred to as a second-price (Vickrey) auction, corresponds to the pivotal mechanism in this setting.

3.7 Problems of the Groves Mechanisms

Lack of Budget Balance in a Public Goods Setting: Consider a situation where there are two individuals and $\Theta_1 = \Theta_2 = \mathbb{R}$. Let the cost of the project be $c = 3/2$.

When $\theta_1 = \theta_2 = 1$, it is efficient to build the public project and hence $x_1(1) + x_2(1) \leq -\frac{1}{2}$. On the hand, $x_1(0) + x_2(0) \leq 0$. Therefore, at least one of $x_1(1) + x_2(0) \leq -\frac{1}{4}$ or $x_1(0) + x_2(1) \leq -\frac{1}{4}$ are satisfied. But since $f(1,0) = f(0,1) = 0$, this implies that the sum of transfers is negative in at least one case.

The above conclusion depends on the richness of the type space. Suppose that in the above example the type spaces only admitted two valuations, $\Theta_1 = \Theta_2 = \{0;1\}$. In that case a simple voting mechanism would induce efficient decisions and no transfers would be necessary. Proposition 23.C.6 in MWG provides a formal statement of this result.

Another problem associated with the Groves mechanisms is the failure of individual rationality or voluntary participation condition. Notice that in the above example, some individual ends up with a negative total utility at (0,1) or (1,0). That individual would have been better off by not participating and obtaining a 0 utility.

4 Bayesian Mechanism Design

The balance difficulties exhibited by Groves' schemes could be overcome in a setting where individuals have probabilistic beliefs over the types of other individuals. This allows us to weaken the requirement of dominant strategy incentive compatibility to a Bayesian incentive compatibility condition.

For simplicity, (as in most of the literature) assume that Θ is a finite set and that $\theta \in \Theta$ is randomly chosen according to a distribution P, where the marginal of P, observes Θ_i, has full support. Each individual knows P and θ_i and has beliefs over the other individuals' types described by Bayes' rule. To distinguish random variables from their realizations,$\bar{\theta}_i$ will denote the random variable and θ_i, θ_i' will denote realizations.

4.1 A Bayesian Revelation Principle

We first define Bayesian equilibrium for a mechanism (M, g). A Bayesian strategy is a mapping $m_i : \Theta_i \to M_i$. A profile of Bayesian strategies $m : \Theta \to M$ forms a Bayesian equilibrium if

$$E\left[v_i(g_d(m_{-i}(\bar\theta_{-i}), m_i(\theta_i)), \theta_i) + g_{t,i}(m_{-i}(\bar\theta_{-i}, m_i(\theta_i)))|\theta_i\right]$$
$$\geq E\left[v_i(g_d(m_{-i}(\bar\theta_{-i}), \hat m_i), \theta_i) + g_{t,i}(m_{-i}(\bar\theta_{-i}, \hat m_i))|\theta_i\right]$$

for each i, $\theta_i \in \Theta_i$ and $\hat m_i \in M_i$. A direct mechanism (i.e., social choice function) $f = (d, t)$ is Bayesian incentive compatible if truth is a Bayesian equilibrium.

The Revelation Principle for Bayesian Equilibrium: If a mechanism (M, g) realizes a social choice function $f = (d, t)$ in Bayesian equilibrium, then the direct mechanism f is Bayesian incentive compatible.

4.2 A Balanced Mechanism

Theorem 3 *If types are independent, d is efficient, and*

$$t_i(\theta) = E\left[\sum_{j\neq i} v_j(d(\bar\theta), \theta_j)|\theta_i\right] - \frac{1}{n-1}\sum_{k\neq i} E\left[\sum_{j\neq k} v_j(d(\bar\theta), \theta_j)|\theta_k\right],$$

then (d, t) is Bayesian incentive compatible and t is balanced.

Proof. The balance of t follows directly from its definition. Let us verify that (d, t) is Bayesian incentive compatible.

$$E\left[v_i(d(\bar{\theta}_{-i},\theta_i'),\theta_i)+t_i(\bar{\theta}_{-i},\theta_i')|\theta_i\right]$$
$$= E\left[v_i(d(\bar{\theta}_{-i},\theta_i'),\theta_i)+\sum_{j\neq i} v_j(d(\bar{\theta}_{-i},\theta_i'),\theta_j)|\theta_i\right] - \frac{1}{n-1}\sum_{k\neq i} E\left[\sum_{j\neq k} v_j(d(\bar{\theta}),\theta_j)|\theta_k\right].$$

The second expression is independent of the announced θ_i'. Since d is efficient, this expression is maximized when $\theta_i' = \theta_i$. ■

The independence condition in Theorem 4 is important in providing the simple structure of the transfer functions, and is critical to the proof. Without independence, it is still possible to find efficient, balanced, Bayesian incentive compatible mechanisms in "most" settings. The extent of "most" has been made precise by d'Aspremont, Cremer, and Gerard-Varet by showing that "most" means except those where the distribution of types is degenerate in that the matrix of conditional probabilities does not have full rank.

Example 3 *(Cremer and McLean (1988)) Consider the Public Project example.* $n = 3$, $c = \frac{7}{4}$ *and* $\theta_i \in \{0,1\}$. *We add a small correlation in the distribution of* θ_i*'s. In particular, we assume that* $\theta = (0,0,0)$ *and* $\theta = (1,1,1)$ *occur with probability* $\frac{1}{8}+3\epsilon$, *while the other realizations of* θ *each occur with probability* $\frac{1}{8}-\epsilon$ *with* $0<\epsilon<\frac{1}{8}$. *So,* ϵ *is a parameter that measures the strength of correlation and adjusts away from the fully independent case. Given the cost, it is efficient to build the project when at least two individuals have* $\theta_i = 1$ *and not otherwise.*

Design a mechanism as follows. Use the efficient decision rule and split costs equally among those with $\theta_i = 1$ *when the project is undertaken. Set* t *as follows.* $t_i(\theta) = x$ *if*

$\theta_i = \theta_{i+1} \neq \theta_{i+2}$; $= -x$ if $\theta_i \neq \theta_{i+1} = \theta_{i+2}$; and $= 0$ otherwise, where $i+1$ and $i+2$ are taken modulo 3. This design utilizes the fact that it is more likely to have $\theta_1 = \theta_2 = \theta_3$. Therefore, if a player misreports, then he is more likely to get punished by paying x.

It is easily seen that t is balanced, and that the mechanism is efficient and interim individually rational. Let us examine the conditions relating x to ϵ that result from requiring Bayesian incentive compatibility. It is easily checked that a type $\theta_i = 0$ does not want to announced $\theta_i = 1$, as that would increase the expected cost paid and decrease the expected transfer. We need only check incentives that type $\theta_i = 1$ not desire to announce $\theta_i = 0$, which requires that

$$(\frac{1}{4} + 6\epsilon)(1 - \frac{c}{3}) + (\frac{1}{4} - 2\epsilon)(1 - \frac{c}{2} + x) + (\frac{1}{4} - 2\epsilon)(1 - \frac{c}{2}) - (\frac{1}{4} - 2\epsilon)x$$
$$\geq (\frac{1}{4} + 6\epsilon)(1 - x) + (\frac{1}{4} - 2\epsilon)0 + (\frac{1}{4} - 2\epsilon)x + (\frac{1}{4} - 2\epsilon)0.$$

This reduces to

$$x \geq \frac{1}{8\epsilon}(\frac{c}{3} - \frac{1}{2} + 6\epsilon),$$

which is satisfied for large enough x.

Note that $x \to \infty$ as $\epsilon \to 0$, which points out one weakness of this approach to exploiting correlation. To take advantage of small amounts of correlation, the size of the transfers has to grow arbitrarily large. If there is some bound on these transfers, or a bankruptcy constraint, then for small amounts of correlation these conditions cannot all be satisfied.

4.3 Bayesian vs. Dominant Strategy Mechanism

Manelli and Vincent (2010) show that in the independent private-values model with linear utility, the outcome - in terms of interim expected probabilities of trade and interim expected transfers - of any Bayesian incentive-compatible mechanism can also be obtained with a dominant-strategy mechanism. In other words, a mechanism is Bayesian incentive compatible if and only if there is a dominant- strategy incentive-compatible mechanism that generates the same interim expected probability of trade for every agent.

5 Optimal Auction Design

5.1 Model Setup

A seller is selling one indivisible object to N risk neutral buyers. Each buyer i's value X_i is distributed over the interval $\mathcal{X}_i = [0, \omega_i]$ according to distribution function F_i. A selling mechanism $(\mathcal{B}, \pi, \mu)$ has the following components: a set of message space $\mathcal{B}_i$ for each buyer i; an allocation rule $\pi : \mathcal{B} \to \Delta(N)$; and a payment rule $\mu : \mathcal{B} \to \mathbb{R}^N$. Each mechanism defines a game of incomplete information among the buyers.

In particular, we will focus on direct mechanism (Q, M) such that $Q : \mathcal{X} \to \Delta(N)$ and $M : \mathcal{X} \to \mathbb{R}^N$. Given a direct mechanism (Q, M), define:

$$q_i(z_i) = \int_{\mathcal{X}_{-i}} Q_i(z_i, x_{-i}) f_{-i}(x_{-i}) dx_{-i}$$

to be the probability that i gets the object. Similarly, define

$$m_i(z_i) = \int_{\mathcal{X}_{-i}} M_i(z_i, x_{-i}) f_{-i}(x_{-i}) dx_{-i}$$

to be the expected payment. The expected payoff of buyer i with true value x_i gets expected payoff $q_i(z_i)x_i - m_i(z_i)$ when he reports z_i. IC constraint implies that

$$U_i(x_i) = q_i(x_i)x_i - m_i(x_i) = \max_{z_i \in \mathcal{X}_i} q_i(z_i)x_i - m_i(z_i).$$

5.2 Characterizing Incentive Compatibility

If a direct mechanism is incentive compatible, then for every agent i, the function q_i is non-decreasing.

If a direct mechanism is incentive compatible, then for every agent i, the function U_i is increasing. It is also convex, and hence differentiable except in at most countably many points. For all i for which it is differentiable, it satisfies: $U_i'(x_i) = q_i$.

Consider an incentive compatible direct mechanism. Then for all i and x_i,

$$U_i(x_i) = U_i(0) + \int_0^{x_i} q_i(t_i) dt_i.$$

Consider an incentive compatible direct mechanism. Then for all i and x_i,

$$m_i(x_i) = m_i(0) + x_i q_i(x_i) - \int_0^{x_i} q_i(t_i) dt_i.$$

5.3 Expected Revenue Maximization

Now we turn to individual rationality. The seller has to choose a mechanism that implies an expected utility of at least zero for the lowest type agents. This is satisfied if and only if $m_i(0) \leq 0$ for all i. If an incentive compatible and individually direct mechanism maximizes the seller's expected revenue then for every i, it must be the case that $m_i(0) = 0$. This allows us to further simplify the seller's expected revenue from any particular buyer i as:

$$\int_{\mathcal{X}_i} \left[x_i q_i(x_i) - \int_0^{x_i} q_i(t_i) dt_i \right] f_i(x_i) dx_i = \int_{\mathcal{X}_i} q_i(x_i) \left[x_i - \frac{1 - F_i(x_i)}{f_i(x_i)} \right] f_i(x_i) dx_i.$$

The total expected revenue from all buyers is:

$$\sum_{i=1}^{N} \int_{\mathcal{X}_i} q_i(x_i) \left[x_i - \frac{1 - F_i(x_i)}{f_i(x_i)} \right] f_i(x_i) dx_i = \sum_{i=1}^{N} \int_{\mathcal{X}} Q_i(x) \left[x_i - \frac{1 - F_i(x_i)}{f_i(x_i)} \right] f(x) dx,$$

where $x = (x_1, \cdots, x_N)$.

Denote $\psi_i(x_i) = x_i - \frac{1-F_i(x_i)}{f_i(x_i)}$ to be buyer i's *virtual value*. Since the above expression is linear in Q_i, it is straightforward to get:

Proposition 1 *(Myerson (1981)) Suppose that for every agent i the cumulative distribution function F_i is regular in the sense that $\psi_i(x_i)$ is strictly increasing. Among all incentive compatible and individually rational direct mechanisms those mechanisms maximize the seller's expected revenue that satisfy for all i and all x: $Q_i(x) = 1$ if $\psi_i(x_i) > 0$ and $\psi_i(x_i) > \psi_j(x_j)$ for all $j \neq i$; and $Q_i(x) = 0$ otherwise.*

The regularity assumption is imposed such that the probability q_i is an increasing function of x_i. Note that we have ignored the case that $\psi_i(x_i) = \psi_j(x_j)$ for some $j \neq i$. This is a zero probability event, and it does not affect either the buyer's incentives or the seller's revenue.

If buyers are symmetric, i.e. the distribution functions F_i are all the same, the optimal mechanism prescribes that the object is given to the buyer with the highest value, if it is sold at all. And the optimal direct mechanism can be implemented using either a first or a second price auction.

However, if buyers are not symmetric, the above rule may not be socially optimal, because social optimality requires the object to be allocated to the buyer with the highest value x_i, while in the optimal mechanism, the object is allocated to the buyer with the highest virtue value $\psi_i(x_i)$.

Economic interpretation of the virtual value: Bulow and Roberts (1989) show that the virtual values are exactly the same as the marginal revenues. Consider any buyer i, and we can reinterpret the cumulative distribution function F_i as a demand curve, where $q = 1 - F_i(v)$ is the demand given price v. Therefore, the marginal revenue can be written as:

$$\frac{dqF_i^{-1}(1-q)}{dq} = v - \frac{1 - F_i(v)}{f_i(v)} = \psi_i(v).$$

The optimal auction derived hence has a very simple interpretation: profit is maximized by maximizing marginal revenue!

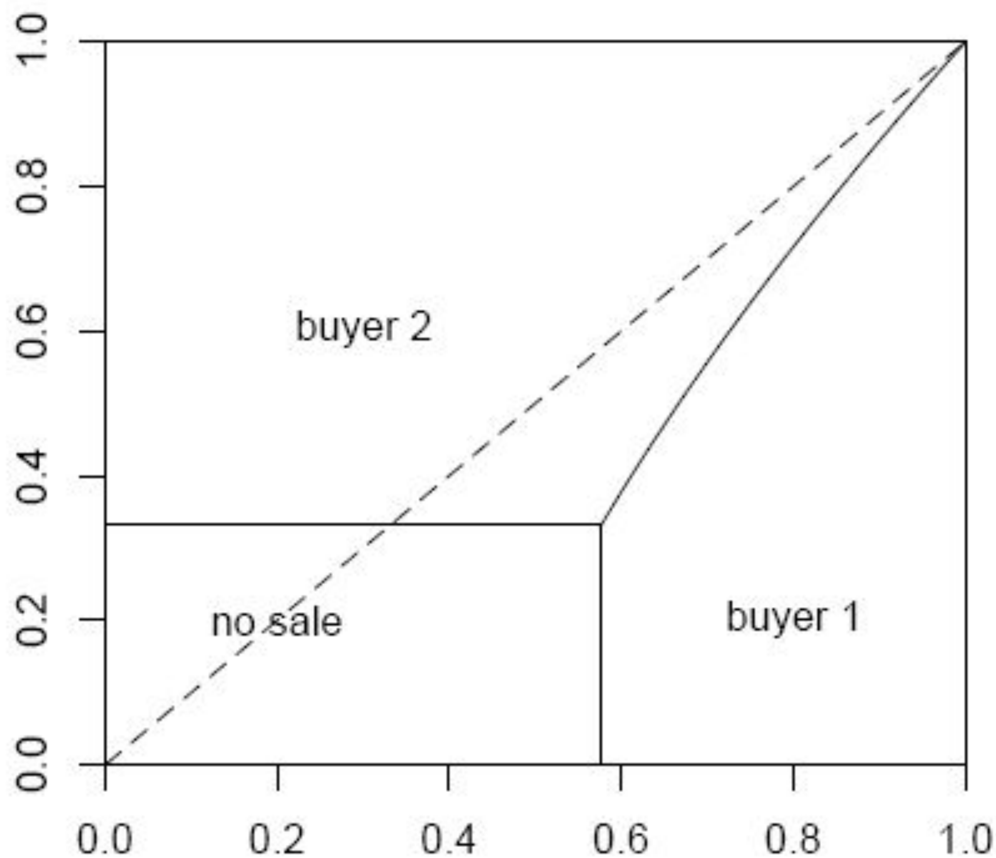
5.4 Numerical Examples

Example 4 *Suppose that* x_i *is uniformly distributed on* $[0, 1]$*. Then it is straightforward to calculate* $\psi(x_i) = 2x_i - 1$*. The regularity condition is satisfied. In the expected revenue maximizing auction the good is sold to neither bidder if* $x_i < \frac{1}{2}$ *for all* i*. The expected revenue maximizing auction will allocate the object to the buyer with the highest value provided that this value is larger than* 0.5*. A first or second price auction with reserve bid* 1/2 *will implement this mechanism. A first or second price auction with reserve bid 0 maximizes expected welfare.*

Example 5 *Suppose that* $N = 2$*, and* $F_1(x_1) = x_1^2$*,* $F_2(x_2) = 2x_2 - x_2^2$ *for* $x_i \in [0, 1]$*. Then it is straightforward to calculate* $\psi_1(x_1) = \frac{3}{2}x_1 - \frac{1}{2x_1}$*,* $\psi_2(x_2) = \frac{3}{2}x_2 - \frac{1}{2}$*. Again, the regularity condition is satisfied. The following Figure shows the optimal allocation of the good. The* 45°*-line is shown as a dashed line. Note that the mechanism is biased against player 1. If the good is sold, bidder 1 wins the object only in a subset of all cases where his value is higher than bidder 2's value. In the expected welfare maximizing mechanism the object is allocated to player 1 if and only if his value is higher than player 2's value.*

5.5 Nonlinear Pricing

Now we study a model in which a monopolist offers an infinitely divisible good, say sugar, to one potential buyer. For simplicity we assume that production costs are linear, that is, producing quantity $q \geq 0$ costs cq, where $c \geq 0$ is a constant. The seller is risk neutral, so that she seeks to maximize her expected revenue. The buyer's utility from buying quantity $q \geq 0$ and paying a monetary transfer t to the monopolist is $\theta v(q) - t$. We assume that


buyer 2
no sale
buyer 1
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.2
0.4
0.6
0.8
1.0

$v(0) = 0$ and that v is a twice differentiable, strictly increasing and strictly concave function.

The parameter θ reflects how much the consumer values the good. The value of θ is known to the buyer but not to the seller. The seller's beliefs about θ are given by a cumulative distribution function F with density f on the interval $[\underline{\theta}, \bar{\theta}]$. We assume that f satisfies: $f(\theta) > 0$ for all $\theta \in [\underline{\theta}, \bar{\theta}]$.

A final assumption is that $\lim_{q \to \infty} \bar{\theta} v'(q) < c$. This means that even the highest type's marginal willingness to pay falls below c as q gets large. This assumption ensures that the quantity that the seller supplies to the buyer is finite for all possible types of the buyer.

We seek to determine optimal selling procedures for the seller. As in the previous section the revelation principle holds and we can restrict attention to direct mechanisms.

A direct mechanism $(q(\theta), t(\theta))$ is incentive compatible if and only if q is increasing and

$$t(\theta) = t(\underline{\theta}) - \underline{\theta} v(q(\underline{\theta})) + \theta v(q(\theta)) - \int_{\underline{\theta}}^{\theta} v(q(\hat{\theta})) d\hat{\theta},$$

for all θ.

An incentive compatible mechanism is individual rational if and only if $t(\underline{\theta}) \leq \underline{\theta} v(q(\underline{\theta}))$. The seller's decision problem is to pick among all direct mechanisms satisfying these two conditions the one that maximizes expected revenue.

Proposition 2 *Suppose that the cumulative distribution function F is regular in the sense that $\psi(\theta) = \theta - \frac{1-F(\theta)}{f(\theta)}$ is strictly increasing. Then an expected profit maximizing choice of q is given by: $q(\theta) = 0$ if $v'(0)\psi(\theta) \leq c$; otherwise, $v'(q(\theta))\psi(\theta) = c$. The profit maximizing t*

is given by:

$$t(\theta) = \theta v(q(\theta)) - \int_{\underline{\theta}}^{\theta} v(q(\hat{\theta}))d\hat{\theta}.$$

A numerical example: $c = 1$, $v(q) = \sqrt{q}$, and $\theta \sim U[0,1]$. It is straightforward to show that $q(\theta) = 0$ if $\theta \leq 0.5$ and $= (\theta - \frac{1}{2})^2$ otherwise. The corresponding $t(\theta) = 0$ if $\theta \leq 0.5$ and $= \frac{1}{2}\theta^2 - \frac{1}{8}$ otherwise. We can translate the solution into an optimal non-linear pricing scheme by expressing the transfer t as a function of q: $t = \frac{1}{2}q + \frac{1}{2}\frac{1}{\sqrt{q}}$. We note that there is a quantity discount. The per unit price decreases in q.

6 Bilateral Trade

6.1 Model Setup

Now we consider the situation from the perspective of a mechanism designer who wants to arrange a trading institution for the seller and the buyer that guarantees that they trade if and only if the buyer's value is larger than the seller's. Not only the buyer's valuation, but also the seller's valuation is unknown to the designer of this trading institution, and it may be that valuations are such that trade is not efficient. This example was first analyzed in the very well-known paper by Myerson and Satterthwaite (1983). It is the simplest example that one might analyze when seeking to build a general theory of the design of optimal trading institution, such as stock exchanges or commodity markets.

A seller S owns a single indivisible good. There is one potential buyer B. The seller's utility if he sells the good and receives a transfer payment t is equal to t. If he does not sell the good and receives a transfer t then his utility is $\theta_S + t$, where θ_S is a random variable with cumulative distribution function F_S and density f_S. We assume that F_S has support $[\underline{\theta}_S, \bar{\theta}_S]$: $f_S(\theta) > 0$ for $\theta \in [\underline{\theta}_S, \bar{\theta}_S]$. The buyer's utility if he purchases the good and pays a transfer t equals $\theta_B - t$, where θ_B is a random variable with cumulative distribution function F_B and density f_B. We assume that F_B has support $[\underline{\theta}_B, \bar{\theta}_B]$. The buyer's utility if he does not obtain the good and pays transfer t is $-t$. The random variables θ_S and θ_B are independent. The seller only observes θ_S, and the buyer only observes θ_B.

6.2 Direct Mechanisms

A direct mechanism consists of functions q, t_S and t_B where: $q : \Theta \to \{0, 1\}$, and $t_i : \Theta \to \mathbb{R}$. The function q assigns to each type vector θ an indicator variable that indicates whether trade takes place ($q = 1$) or whether no trade takes place ($q = 0$). For simplicity, we restrict attention to deterministic trading rules. The function t_S indicates transfers that the seller receives, and the function t_B indicates transfers that the buyer makes.

Given a direct mechanism, we define for each agent $i \in \{S, B\}$ functions $Q_i : [\underline{\theta}_i, \bar{\theta}_i] \to [0, 1]$ and $T_i : [\underline{\theta}_i, \bar{\theta}_i] \to \mathbb{R}$, where Q_i is the conditional probability that trade takes place, conditioning on agent i's type being θ_i; and T_i is the conditional expected value of the transfer that agent i receives (if $i = S$) or makes (if $i = B$), again conditioning on agent i's type being θ_i. Finally, we also define agent i's expected utility $U_i(\theta_i)$ conditional on her type being θ_i. This is given by: $U_S(\theta_S) = T_S(\theta_S) + (1 - Q_S(\theta_S))\theta_S$ and $U_B(\theta_B) = -T_B(\theta_B) + Q_B(\theta_B)\theta_B$.

We shall restrict attention to direct mechanisms that are incentive compatible, individually rational, and ex post budget balanced. Individual rationality is defined as before. Standard arguments show that a mechanism is incentive compatible for the buyer under exactly the same conditions as before. For the seller, the standard arguments apply if types are ordered in the reverse of the numerical order, that is starting with high types rather than low types. Thus, a necessary and sufficient condition for incentive compatibility for the seller is that Q_S is decreasing, and that T_S is given by:

$$T_S(\theta_S) = T_S(\bar{\theta}_S) + (1 - Q_S(\bar{\theta}_S))\bar{\theta}_S - (1 - Q_S(\theta_S))\theta_S - \int_{\theta_S}^{\bar{\theta}_S} (1 - Q_S(x))dx$$

for all θ_S.

Individual rationality for the buyer is defined and characterized in the same way as before. For the seller, individual rationality means that $U_S(\theta_S) \geq \theta_S$ for all θ_S: the seller trades voluntarily and obtains an expected utility that is at least as large as his utility would be if he kept the good. If a mechanism is incentive compatible, then the seller's individual rationality condition holds if and only if it holds for the highest seller type $\bar{\theta}_S$.

Ex post budget balance requires that in each state θ we have: $t_S(\theta) = t_B(\theta)$

6.3 Welfare Maximization

The mechanism designer seeks to maximize the sum of the individuals' utilities,

$$\theta_S + q(\theta)(\theta_B - \theta_S) + t_S - t_B.$$

If the mechanism designer were not constrained by incentive compatibility and individual rationality, but only had to respect ex post budget balance, then the mechanism designer would choose a first best mechanism where the trading rule is: $q^\star(\theta) = 1$ if $\theta_B \geq \theta_S$; and

$= 0$ otherwise.

Proposition 3 *Myerson and Satterthwaite (1983). An incentive compatible, individually rational and ex-post budget balanced direct mechanism with decision rule* $q^\star$ *exists if and only if* $\underline{\theta}_B \geq \bar{\theta}_S$ *or* $\underline{\theta}_S \geq \bar{\theta}_B$*.*

The condition $\underline{\theta}_B \geq \bar{\theta}_S$ implies that trade is always at least weakly efficient. The condition $\underline{\theta}_S \geq \bar{\theta}_B$ implies that efficiency never requires trade. Thus, these are trivial cases. In all non-trivial cases, there is no incentive compatible, individually rational and ex post budget balanced first best mechanism.

Proof. The "if-part" is trivial. If $\underline{\theta}_S \geq \bar{\theta}_B$, then a mechanism under which no trade takes place and no payments are made is first best and has the required properties. If $\underline{\theta}_B \geq \bar{\theta}_S$, a mechanism where trade always takes place, and the buyer always pays the seller some price $p \in [\underline{\theta}_S, \bar{\theta}_B]$ is first best and has the required properties.

To prove the "only if-part", we proceed in two steps as in MWG. The first step shows that:

$$\int_{\theta \in \Theta} q(\theta) \left[\psi_B(\theta_B) - \psi_S(\theta_S)\right] f(\theta) d\theta \geq 0,$$

where $\psi_B(\theta_B) = \theta_B - \frac{1-F_B(\theta_B)}{f_B(\theta_B)}$ and $\psi_S(\theta_S) = \theta_S + \frac{F_S(\theta_S)}{f_S(\theta_S)}$. The second step shows that the above inequality cannot be satisfied under $q^\star$. ■

Intuitively, in any two-sided negotiation between a buyer and seller, the seller has an incentive to exaggerate its value and the buyer has an incentive to pretend its value is lower. These misrepresentations imply that there is no way to design a bargaining protocol that avoids this problem: delays or failures are inevitable in private bargaining if the good starts out in the wrong hands.

6.4 A Numerical Example

Example 6 *Suppose that both* θ_B *and* θ_S *are uniformly distributed on the interval* $[0,1]$*. We want to determine the welfare maximizing incentive compatible, individually rational and ex post budget balanced trading mechanism.*

Proposition 4 *In the welfare maximizing incentive compatible, individually rational and ex post budget balanced trading mechanism trade takes place if and only if* $\theta_B - \theta_S \geq \frac{1}{4}$.

Proof. *The mechanism designer seeks to maximize:*

$$\int_{\theta\in\Theta} q(\theta)(\theta_B - \theta_s) f(\theta) d\theta,$$

subject to the following constraints:

1. Q_B *is increasing and* Q_S *is decreasing;*

2. $U_S(\bar{\theta}_s) \geq \bar{\theta}_S$ *and* $U_B(\underline{\theta}_B) \geq 0$*;*

3. $$-U_B(\underline{\theta}_B) + \int_{\theta\in\Theta} q(\theta)\psi_B(\theta_B) f(\theta) d\theta = U_S(\bar{\theta}_S) - \int_{\theta\in\Theta} (1 - q(\theta))\psi_S(\theta_S) f(\theta) d\theta.$$

Plug constraint 2 into 3, and we can get:

$$\int_{\theta\in\Theta} q(\theta)\left[\psi_B(\theta_B) - \psi_S(\theta_S)\right] f(\theta) d\theta \geq K = \bar{\theta}_S - \int_{\theta\in\Theta} \psi_S(\theta_S) f(\theta) d\theta.$$

It is straightforward to show that $K = 0$.

A necessary and sufficient condition for a mechanism q to be optimal is that there is a Lagrange multiplier $\lambda \geq 0$ such that q maximizes:

$$\int_{\theta\in\Theta} q(\theta)\left[\theta_B - \theta_S + \lambda(\psi_B(\theta_B) - \psi_S(\theta_S))\right] f(\theta) d\theta.$$

Therefore, $q(\theta) = 1$ if $\theta_B - \theta_S + \lambda(\psi_B(\theta_B) - \psi_S(\theta_S)) \geq 0$; and $= 0$ otherwise. Since both θ_B and θ_S are uniformly distributed on the interval $[0, 1]$, the above condition is simplified as:

$$\theta_B - \theta_S \geq s = \frac{\lambda}{1 + 2\lambda}.$$

Take s as given. The buyer's expected payment is

$$\int_{\theta\in\Theta} q(\theta)\psi_B(\theta_B) f(\theta) d\theta = \frac{1}{3}s^3 - \frac{1}{2}s^2 + \frac{1}{6}.$$

The seller's expected transfer is

$$\bar{\theta}_S - \int_{\theta\in\Theta} (1 - q(\theta))\psi_S(\theta_S) f(\theta) d\theta = \frac{1}{3}(1 - s)^3.$$

Budget balance is achieved if $\frac{1}{3}s^3 - \frac{1}{2}s^2 + \frac{1}{6} = \frac{1}{3}(1 - s)^3$, which implies that $s = \frac{1}{4}$. ■

The optimal mechanism can be implemented using a double auction: Seller and buyer simultaneously propose prices $p_s \in [0, 1]$ and $p_b \in [0, 1]$. Trade occurs at price $\frac{1}{2}(p_b + p_s)$ if $p_b \geq p_s$; otherwise no trade.

6.5 Overcoming Participation Constraints

The above example is widely viewed as a counter example of the well-known "Coase Theorem". Consider two players with uniformly distributed valuations on the interval $[0, 1]$. Efficiency can be achieved by using a first/second price auction. However, if the property rights of the good is allocated to a player, then efficiency cannot be achieved. The key insights are the allocation of the good affects the owner's incentive constraints, which limits the attainment of the socially efficient outcomes. Now the question is how to overcome these participation constraints.

6.5.1 Cramton, Gibbons, and Klemperer (1987)

Cramton, Gibbons, and Klemperer (1987) consider the case in which property rights are divisible, e.g., the asset is perfectly divisible with agent i's property right providing him with payoff $\theta_i x_i$ when he owns fraction $x_i \in [0, 1]$. They show that efficient bargaining is possible (i.e., there is an efficient, incentive-compatible, individually rational, and budget-balanced mechanism) provided that the agents start from an allocation of property rights that is close enough to equal. Intuitively, intermediate ownership levels reduce the incentives for misrepresentation because each agent does not know whether he will ultimately end up as a seller or a buyer of the asset.

Consider the example with two players and independent uniform distribution. Efficient trading mechanism can be achieved by the following trading game: each player has 0.5 share of the good, and they submit sealed bids. The good is transfered to the highest bidder, and the price paid by player i is $p_i = b_i - b_j$.

It is straightforward to show that in this bidding game, there is a symmetric BNE: $b(\theta) = \frac{1}{2}\theta^2$. This equilibrium is obviously efficient because player with higher θ will obtain the good.

The last thing is to verify individual rationality. Will each play with type θ has incentive to engage in this bidding game? By not participating, the payoff is $\frac{1}{2}\theta$. By participating, the expected payoff is:

$$\int_0^\theta (\theta + \frac{1}{2}x^2)dx + \int_\theta^1 \frac{1}{2}x^2 dx - \frac{1}{2}\theta^2 = \frac{1}{2}\theta^2 + \frac{1}{6} > \frac{1}{2}\theta.$$

If the allocation of property rights is close enough to equal, Cramton, Gibbons, and Klemperer (1987) show that efficiency can be achieved in the above bidding game by adding side payments depending on the allocation of property rights. See Segal and Whinston (2011) for a recent development of the theory.

6.5.2 Jackson and Sonnenschein (2007)

Jackson and Sonnenschein (2007) propose another way to overcome incentive constraints by linking decisions, based on the original idea of Cohn (2010). Consider the following bargaining problem. There is a seller of an object with valuation chosen uniformly from $\{0.1, 0.3, 0.5, 0.7, 0.9\}$ and a buyer with valuation chosen uniformly from $\{0.2, 0.4, 0.6, 0.8, 1\}$.

Approximately efficiency can be achieved by linking K (say divisible by 5) decision problems by requiring each agent to specify exactly 1/5 of the problems where they have each valuation. For example, if $K = 50$, the seller has to announce valuation 0.1 on 10 problems, 0.3 on 10 problems, etc. Similarly, the buyer has to announce valuation 0.2 on 10 problems, 0.4 on 10 problems, etc. On each problem, the agents trade the object if and only if the buyer's value exceeds that of the seller, and the price is the average of the valuations. There is an approximately truthful equilibrium where agents tell the truth to the maximal extent possible, given that it is possible that they will not have a given valuation on exactly 1/5 of the problems. For large K, the fraction of problems where the correct decision is made goes to 1 in probability.

References

Bulow, J., and J. Roberts (1989): "The Simple Economics of Optimal Auctions," *Journal of Political Economy*, 97, 1060–1090.

Cohn, Z. (2010): "A Note on Linked Bargaining," *Journal of Mathematical Economics*, 46, 238–247.

Cramton, P., R. Gibbons, and P. Klemperer (1987): "Dissolving a Partnership Efficiently," *Econometrica*, 55, 615–32.

Cremer, J., and R. P. McLean (1988): "Full Extraction of the Surplus in Bayesian and Dominant Strategy Auctions," *Econometrica*, 56, 1247–1258.

Jackson, M., and H. Sonnenschein (2007): "Overcoming Incentive Constraints by Linking Decisions," *Econometrica*, 75(1), 241–257.

Manelli, A. M., and D. R. Vincent (2010): "Bayesian and Dominant-Strategy Implementation in the Independent Private Values Model," *Econometrica*, 78, 1905–1938.

Myerson, R. (1981): "Optimal Auction Design," *Mathematics of Operations Research*, 6, 58–73.

Myerson, R., and M. Satterthwaite (1983): "Efficient Mechanisms for Bilateral Trading," *Journal of Economic Theory*, 29, 265–81.

Segal, I., and M. D. Whinston (2011): "A simple status quo that ensures participation (with application to efficient bargaining)," *Theoretical Economics*, 6(1), 109–125.