

Foundations of mechanism design: A tutorial Part 1 – Key concepts and classical results

DINESH GARG¹, Y NARAHARI^{2*} and SUJIT GUJAR²

¹IBM India Research Laboratory, Bangalore 560 071

²Electronic Commerce Laboratory, Department of Computer Science and Automation, Indian Institute of Science, Bangalore 560 012
e-mail: dingarg2@in.ibm.com; {hari,sujit}@csa.iisc.ernet.in

MS received 6 November 2006; revised 16 December 2007

Abstract. Mechanism design, an important tool in microeconomics, has found widespread applications in modelling and solving decentralized design problems in many branches of engineering, notably computer science, electronic commerce, and network economics. Mechanism design is concerned with settings where a social planner faces the problem of aggregating the announced preferences of multiple agents into a collective decision when the agents exhibit strategic behaviour. The objective of this paper is to provide a tutorial introduction to the foundations and key results in mechanism design theory. The paper is in two parts. Part 1 focuses on basic concepts and classical results which form the foundation of mechanism design theory. Part 2 presents key advanced concepts and deeper results in mechanism design.

Keywords. Mechanism design; game theory; social choice functions; auctions.

1. Introduction

In the second half of the twentieth century, game theory and mechanism design have found widespread use in a gamut of applications in engineering. More recently, game theory and mechanism design have emerged as an important tool to model, analyse, and solve *decentralized design problems* in engineering involving multiple autonomous agents that interact strategically in a *rational* and *intelligent* way. The importance of mechanism design in the current context can be seen by the fact that the Nobel Prize in Economic Sciences for the year 2007 was jointly awarded to three economists, Leonid Hurwicz, Eric Maskin, and Roger Myerson *for having laid the foundations of mechanism design theory* (The Nobel Foundation 2007). Earlier, in 1996, William Vickrey, the inventor of the famous Vickrey auction had been awarded the Nobel Prize in Economic Sciences.

The agents are *rational* in the game theoretic sense of making decisions consistently in pursuit of their own individual objectives. Each agent's objective is to maximize the

*For correspondence

expected value of his/her own payoff measured in some utility scale. *Selfishness* or *self-interest* is an important implication of rationality. Each agent is *intelligent* in the game theoretic sense of knowing everything about the underlying game that a game theorist knows and each agent can make any inferences about the game that a game theorist can make. In particular, each agent is *strategic*, that is, takes into account his/her knowledge or expectation of behaviour of other agents and is capable of doing the required computations.

The theory of *mechanism design* is concerned with settings where a policy maker (or social planner) faces the problem of aggregating the *announced preferences* of multiple agents into a collective (or social) decision when the *actual preferences* are not publicly known. Mechanism design theory uses the framework of *non-cooperative games with incomplete information* and seeks to study how the privately held preference information can be elicited and the extent to which the information elicitation problem constrains the way in which social decisions can respond to individual preferences. In fact, mechanism design can be viewed as *reverse engineering* of games or equivalently as the *art of designing the rules of a game to achieve a specific desired outcome*. The main focus of mechanism design is to design institutions or protocols that satisfy certain desired objectives, assuming that the individual agents, interacting through the institution, will act strategically and may hold private information that is relevant to the decision at hand.

1.1 Mechanism design: A brief history

Leonid Hurwicz (Nobel laureate in Economic Sciences in 2007) first introduced the notion of mechanisms with his work in 1960 (Hurwicz 1960). He defined a mechanism as a communication system in which participants send messages to each other and perhaps to a *message center* and a pre-specified rule assigns an outcome (such as allocation of goods and payments to be made) for every collection of received messages. William Vickrey (1961) (Nobel laureate in Economic Sciences in 1996) wrote a classic paper (Vickrey 1961) which introduced the famous Vickrey auction (second price auction). To this day, the Vickrey auction continues to enjoy a special place in the annals of mechanism design. John Harsanyi (1967, 1968a, 1968b) (Nobel laureate in Economic Sciences in 1994 jointly with John Nash and Richard Selten) developed the theory of games with incomplete information, in particular Bayesian games, through a series of three seminal papers. Harsanyi's work later proved to be of foundational value to mechanism design. Hurwicz (1972) introduced the key notion of incentive compatibility. This notion allowed mechanism design to incorporate the incentives of rational players and opened up mechanism design. Clarke (1971) and Groves (1973) came up with a generalization of the Vickrey mechanisms and helped define broad class of dominant strategy incentive compatible mechanisms in the quasi-linear environment.

There were two major advances in mechanism design in the 1970s. The first was the *revelation principle* which essentially showed that direct mechanisms are the same as indirect mechanisms. This meant that mechanism theorists needed to worry only about direct mechanisms, leaving the development of real-world mechanisms (which are mostly indirect mechanisms) to mechanism designers and practitioners. Gibbard (1973) formulated the revelation principle for dominant strategy incentive compatible mechanisms. This was later extended to Bayesian incentive compatible mechanisms through several independent efforts (The Nobel Foundation 2007) — Maskin and Myerson (both Nobel laureates in Economic Sciences in 2007) had a leading role to play in this. In fact, Myerson developed the revelation principle in its greatest generality (The Nobel Foundation 2007). The second major advance in mechanism design in the 1970s was on *implementation theory* which addresses the following

problem: Can a mechanism be designed so that all its equilibria are optimal? Maskin (1999) gave the first general solution to this problem.

Mechanism design has made numerous significant advances during 1980s, 1990s, and during the past few years. It has found widespread applicability in a variety of disciplines. These include: design of markets and trading institutions (The Nobel Foundation 2007, Mas-Colell *et al* 1995, Milgrom 2004), regulation and auditing (The Nobel Foundation 2007), social choice theory (The Nobel Foundation 2007), computer science (Nisan *et al* 2007), and network economics (Narahari *et al* 2008). The above list is by no means exhaustive. In this tutorial, our discussion is motivated by applications of mechanism design in network economics which is concerned with an economics based approach to solving design problems in distributed systems and networks. We now provide below several examples from network economics.

1.2 Examples of mechanism design problems in network economics

1.2a Mechanism design and algorithms for selfish agents: In distributed settings, there are certain algorithmic problems where the agents cannot be assumed to follow the algorithm but are driven by selfish goals (Nisan 1999). In such situations, the agents are capable of manipulating the algorithm. For example, in a shortest path problem where edges are owned by individual agents and the costs of the edges are known only to the owning agents, the algorithm designer is faced with the challenge of eliciting the true costs of the edges from the owning agents before applying an algorithm for computing the shortest paths. The objective of the algorithm designer in such situations should be to come up with a scheme which ensures that the agents' interests are best served by behaving truthfully. Mechanism design theory provides a handle for studying and designing such algorithms. The mechanism is designed in a way that all agents are motivated to act truthfully and according to the wishes of the algorithm designer.

1.2b Mechanism design and selfish routing: A major difficulty that is faced by networks such as road transportation networks, communication networks, and the Internet is that most of the time, the demand of the users for the network resources exceeds the available supply of the resources. This phenomenon causes congestion in the network. The traffic congestion can be avoided or mitigated if the arrival rates can be controlled and/or the traffic can be routed across the network in an effective manner. However, in most of these networks, the users of such networks are free to act according to their own interests and moreover, they are rational and intelligent in the sense that they care more for their own individual welfare and less for the health or performance of the overall network. Game theory and mechanism design can play an important role in the analysis and design of protocols in such networks. Indeed, in the past few years, there has been a spurt of research activities in this direction. See, for example, the work of Roughgarden (2005), Feigenbaum *et al* (2002), Feigenbaum & Shenker (2002), Hershberger & Suri (2001), and Nisan & Rohen (2000, 2001).

1.2c Mechanism design and ad-hoc networks: Wireless ad-hoc networks also face the traffic congestion problem just like other wired networks such as communication networks and the Internet. However, due to the wireless nature of the communication medium, the physics of the problem is somewhat different. In the case of wireless ad-hoc networks, conservation of battery power by the individual nodes is of primary importance. Individual nodes are required to forward packets so as to ensure connectivity in an ad-hoc network and nodes therefore double up as routers. Forwarding of packets involves draining battery power and it

may not always be in the self-interest of a node to forward packets. Design of protocols that stimulate cooperative actions from the nodes uses game theory and mechanism design in a significant way. There are several recent efforts in this direction. See, for example, the work of Anderegge & Eidenbenz (2003) in unicast environment, where they use the VCG (Vickrey–Clarke–Groves) mechanism to compute a power efficient path using which the nodes transmit or forward the packets. Eidenbenz, *et al* (2005) further generalize this model to achieve budget balance property. Wang *et al* (2004) propose strategy-proof pricing schemes for multicast environment. Suri (2006) has generalized the model of Wang *et al* (2004) and has proposed incentive compatible broadcast protocols in ad-hoc wireless networks.

1.2d Mechanism design and grid computing: One of the most important issues concerned with grid computing is that of application scheduling. In a global grid setting, the individual users must be provided with an incentive to offer their resources. The situation becomes non-trivial because of the fact that these entities are rational and intelligent resource providers who, for strategic reasons, may not provide truthful information about their processing power and cost structure. Thus, resource allocation decisions in computational grids and other general grids have to take into account the rationality of the grid resource providers. In particular, there is a need for providing appropriate incentives to the nodes to stimulate their participation in grid computing. Mechanism design is useful in designing incentives and also for auction based mechanisms for grid resource allocation. There are several recent efforts in this direction. See, for example, the work of Grosu & Chronopoulos (2003, 2004), Das & Grosu (2005), Walsh *et al* (1998), and Buyya (2002). Prakash & Narahari (2005) have recently proposed several innovative mechanisms for resource procurement in computational grids with rational resource providers.

1.2e Mechanism design and cryptography: Mechanism design can be viewed as a science of synthesis of protocols for selfish parties to achieve certain properties. In mechanism design, it is often assumed that the central authority (or social planner) can be trusted by the parties, but this might not always be true, especially in an Internet environment. If the social planner is corrupt, then it may misuse the information received from the agents. Privacy is therefore essential in order to ensure the social planner's credibility (Naor 2001). This problem was first stated by Varian (1995). Because cryptography deals with preserving privacy and integrity of data in computer and communication systems, it is obvious that techniques from cryptography may help implementing the mechanism design protocols in the real world. There has been some progress in the recent past in applying cryptographic tools and techniques to the problems of mechanism design. For example, see the work related to privacy preserving auctions by Noar *et al* (1999) and Brandt (2003).

1.2f Mechanism design and the world wide web: We know that the world wide web has become an integral part of our day-to-day life. As engineers and scientists are making progress in developing innovative web based services, the hackers and spammers are bent on making such services collapse. Examples of such services include *search engines page ranking system*, *recommender systems*, and *reputation systems*. Note that it is not always the case that these hackers and spammers invest their energy and efforts just for nothing. More often than not, they do have a stake in not letting such systems run smoothly. The only solution for such a problem is to design the system in a way that there is no incentive for the hackers and spammers to put their energy and time in destroying it. Game theory and mechanism design theory play an important role in designing such systems. For example see the work of

Altman & Tennenholtz (2005, 2006) for a mechanism design view of page ranking systems, and Gyongyi & Garcia-Molina (2005) for a kind of threat to the page ranking system, namely link spam.

1.2g *Mechanism design and electronic commerce*: Electronic commerce is an area where the rationality in general and self-interest in particular of the participating agents is a significant factor to be taken into account. Game theory and mechanism design, therefore, have come to play a natural role here. For example, in a bargaining problem between a buyer and a seller, the seller would like to act as if the item is very expensive thus raising its price and the buyer would like to pretend to have a low value for the object to keep the price down. In this context, mechanism design helps to come up with a bargaining protocol that ensures an efficient trade of the good, so that successful trade occurs whenever the buyer's valuation exceeds that of the seller.

There are a variety of game theoretic design problems currently being explored in the area of electronic commerce. These design problems are well-documented in the literature; the following is a representative listing:

- Auctions for selling spectrum (Cramton 2005)
- Mechanisms for selling advertising space through keyword auctions (Edelman *et al* 2006, Garg 2006, Aggarwal *et al* 2006)
- Auctions for selling products/services as part of private marketplaces set-up by e-business companies (Narahari & Dayama 2005)
- Bandwidth exchanges (Eso *et al* 2001)
- Procurement auctions and private marketplaces for e-procurement (Chandrashekar *et al* 2006, Kameshwaran *et al* 2006, Gautam *et al* 2007)
- Logistics and transportation marketplaces (Caplice & Sheffi 2005)
- Mechanisms for supply chain formation (Walsh *et al* 2000, Walsh & Wellman 2003, Karabuk & David Wu 2005, Narahari & Srivastava 2007).

1.3 Outline of the paper

The focus of this Part 1 is on basic concepts and key results that form the foundation of mechanism design theory. We organize the material in this part as follows.

- We start by introducing the notion of a *social choice function* in § 2, through several examples: bilateral trade, auctions for selling a single indivisible item, and a combinatorial auction. These examples are used throughout the rest of the paper for bringing out important insights. In § 3, we introduce the concept of a *mechanism* and bring out the difference between direct revelation mechanisms and indirect mechanisms. Section 4 is devoted to the notion of *implementation* of a social choice function by a mechanism. We introduce two key notions of implementation, namely dominant strategy implementation and Bayesian implementation.
- Next, we describe, in § 5, the desirable properties of a social choice function, which include *ex-post efficiency*, *non-dictatorialness*, *dominant strategy incentive compatibility* and *Bayesian–Nash incentive compatibility*. We then state and prove a fundamental result in mechanism design theory, the *revelation theorem* in § 6.
- Next we introduce, in § 7, a landmark result — the *Gibbard–Satterthwaite* impossibility theorem, which says that under fairly general conditions, no social choice function can satisfy the three properties — *ex-post efficiency*, *non-dictatorial*, and *dominant strategy*

incentive compatibility simultaneously. This impossibility theorem is a special case of the celebrated Arrow's impossibility theorem, which we present in § 8.

- The Gibbard–Satterthwaite theorem, while ruling out implementability of certain desirable mechanisms, suggests two alternative routes to design useful mechanisms. The first of these two routes is to restrict the utility functions to what is known as a *quasi-linear environment*. We introduce the quasi-linear environment in § 9, where we show that ex-post efficiency is equivalent to a combination of two properties, namely, *allocative efficiency* and *budget balance*. In § 10, we study the celebrated VCG (*Vickrey–Clarke–Groves*) social choice functions (also known as VCG mechanisms), which are non-dictatorial, dominant strategy incentive compatible, and allocatively efficient.
- We explore the second route suggested by the Gibbard–Satterthwaite impossibility theorem in § 11. Here we look for Bayesian incentive compatibility instead of dominant strategy incentive compatibility. We show that a class of social choice functions, known as *dAGVA* social choice functions are ex-post efficient, non-dictatorial, and Bayesian–Nash incentive compatible. We then develop a characterization of Bayesian incentive compatible social choice functions in linear environment.
- Finally, in § 12, we provide pointers to further literature on all aspects discussed in this part.

In Part 2 of this tutorial (Garg *et al* 2008), we build upon the basic concepts and key results discussed in Part 1 and look into deeper issues and results in mechanism design theory. The topics that we cover in Part 2 include: (1) Revenue equivalence of auctions; (2) Individual rationality; (3) Moulin mechanisms; (4) Optimal auctions; (5) Characterization of dominant strategy incentive compatible (DSIC) mechanisms; (6) DSIC implementation of Bayesian incentive compatible (BIC) rules; (7) Implementation in ex-post Nash equilibrium; (8) Mechanisms with interdependent types; (9) Implementation of mechanisms; and (10) Other advanced topics in mechanism design.

This tutorial assumes familiarity with basic game theory, in particular topics such as strategic form games, Bayesian games, and notions such as dominant strategy equilibria, Nash equilibria, and Bayesian–Nash equilibria. There are numerous texts available, for example, the ones by Martin (2003) and Myerson (1997).

2. Social choice function

DEFINITION 2.1.

A social choice function (SCF) is a function $f: \Theta \rightarrow X$, which a social planner or policy maker uses to assign a collective choice $f(\theta_1, \dots, \theta_n)$ to each possible profile of the agents' types $\theta = (\theta_1, \dots, \theta_n) \in \Theta$.

Figure 1 illustrates the idea behind social choice function. In what follows, we present a few examples of the mechanism design problem and the social choice function that is being used by the social planner. Most of these examples are taken from (Mas-Colell *et al* 1995).

2.1 Example: Allocation of a single unit of an indivisible private good

Consider a situation where there is a set N of n agents and one of them is owning one unit of an indivisible good. The owner wants to trade this good with other agents by means of money. The other agents are interested in trading with him. Let us assume that there is an outside

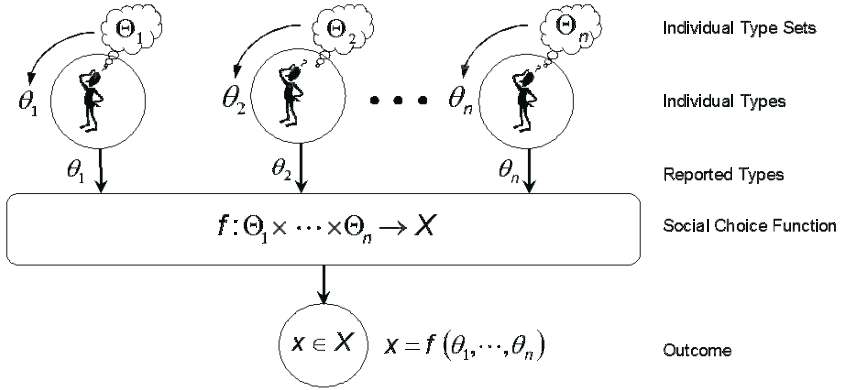


Figure 1. The idea behind social choice function.

agent who plays the role of a broker and facilitates the trading activities among these agents. Each of the n agents interact directly with the broker agent in a personal manner but they do not interact among themselves. This broker agent can be viewed as a social planner. Here the problem of the broker agent is to decide whom to allocate the good and how much money to charge (or pay) from each agent. This can be viewed as a mechanism design problem and the various building blocks of the underlying mechanism design structure can be given in following manner.

2.1a *Outcome set X* : An outcome in this case may be represented by a vector $x = (y_1, \dots, y_n, t_1, \dots, t_n)$, where $y_i = 1$ if the agent i receives the object, $y_i = 0$ otherwise, and t_i is the monetary transfer received by the agent i . The set of feasible alternatives is then

$$X = \left\{ (y_1, \dots, y_n, t_1, \dots, t_n) \mid y_i \in \{0, 1\}, t_i \in \mathbb{R} \ \forall i, \sum_{i=1}^n y_i = 1, \sum_{i=1}^n t_i \leq 0 \right\}.$$

2.1b *Type set Θ_i* : In this example, the type θ_i of an agent i can be viewed as his valuation of the good. We can take the set of possible valuations for agent i to be $\Theta_i = [\underline{\theta}_i, \bar{\theta}_i] \subset \mathbb{R}$, where $\underline{\theta}_i$ is the least valuation and $\bar{\theta}_i$ is the highest valuation the agent may have for the item.

2.1c *Utility function $u_i(\cdot)$* : The utility function of agent i can be given by

$$\begin{aligned} u_i(x, \theta_i) &= u_i(y_1, \dots, y_n, t_1, \dots, t_n, \theta_i) \\ &= \theta_i y_i + t_i. \end{aligned}$$

2.1d *Social choice function $f(\cdot)$* : The general structure of the social choice function for this case is

$$f(\theta) = (y_1(\theta), \dots, y_n(\theta), t_1(\theta), \dots, t_n(\theta)) \ \forall \theta \in \Theta.$$

Two special cases of this example have received a great deal of attention in the literature — *bilateral trading* and *auctions*. In what follows, we discuss each of these two.

2.2 Example: Bilateral trade

This is a special case of the previous example. In this setting, we have $n = 2$. Agent 1 is interpreted as the initial owner of the good (the ‘seller’), and agent 2 is the potential purchaser of the good (the ‘buyer’). A few facts are in order. Note that

- (i) If $\overline{\theta}_1 < \underline{\theta}_2$ there are certain gains for both the agents from the trade regardless of θ_1 and θ_2 .
- (ii) If $\overline{\theta}_2 < \underline{\theta}_1$ then it is certain there are no gains for both the agents from the trade.
- (iii) If $\underline{\theta}_2 < \overline{\theta}_1$ or $\underline{\theta}_1 < \overline{\theta}_2$ or both, then there may or may not be gains for the agents from trade, depending on the realization of θ .

We will develop this example further in this paper as and when required.

2.3 Example: Single unit–single item auction

This is a special case of the example 2.1. For the purpose of illustration, we assume that there are $n + 1$ agents instead of n agents. Note that there is no loss of generality in doing so. Let agent 0 be the owner of the good with the rest of the n agents interested in buying the good. In this case, the owner of the good is known as auctioneer and the rest of the agents are known as bidders. The distinguishing feature of this example which makes it a special case of the Example 2.1 is the following:

The type set of the seller is a singleton set, that is $\Theta_0 = \{\theta_0\}$, which is commonly known to all the buyers.

If $\theta_0 = 0$ then it implies that the auctioneer has no value for the good and he just wants to sell it off. We call such a scenario as single unit single item auction without reserve price. For this example, the various building blocks of the underlying mechanism design structure are the following.

$$X = \left\{ (y_0, y_1, \dots, y_n, t_0, t_1, \dots, t_n) \mid y_0 = 0, t_0 \geq 0, y_i \in \{0, 1\}, \right.$$

$$\left. t_i \leq 0 \forall i = 1, \dots, n, \sum_{i=0}^n y_i = 1, \sum_{i=0}^n t_i = 0 \right\}$$

$$\Theta_0 = \{\theta_0 = 0\}$$

$$\Theta_i = [\underline{\theta}_i, \overline{\theta}_i] \subset \mathbb{R} \forall i = 1, \dots, n$$

$$u_0(x, \theta_0) = u_0(y_0, y_1, \dots, y_n, t_0, t_1, \dots, t_n, \theta_0) = t_0$$

$$u_i(x, \theta_i) = u_i(y_0, y_1, \dots, y_n, t_0, t_1, \dots, t_n, \theta_i) = \theta_i y_i + t_i \quad \forall i = 1, \dots, n$$

$$f(\theta) = (y_0(\theta), \dots, y_n(\theta), t_0(\theta), \dots, t_n(\theta)) \quad \forall \theta \in \Theta.$$

2.4 Example: Single unit–single item auction with reserve price

This is the same example as the previous one except that now auctioneer has some positive value for the good, that is, $\theta_0 > 0$. The auctioneer announces a reserve price $r > 0$, which

need not be the same as θ_0 . All the other agents treat this reserve price as the valuation of the auctioneer for the good. This scenario is known as single unit–single item auction with reserve price. For this example, the various building blocks of the underlying mechanism design structure are the following.

$$X = \left\{ (y_0, y_1, \dots, y_n, t_0, t_1, \dots, t_n) \mid y_0 \in \{0, 1\}, t_0 \geq 0, y_i \in \{0, 1\}, \right.$$

$$\left. t_i \leq 0 \forall i = 1, \dots, n, \sum_{i=0}^n y_i = 1, \sum_{i=0}^n t_i = 0 \right\}$$

$$\Theta_0 = \{\theta_0 > 0\}$$

$$\Theta_i = [\underline{\theta}_i, \bar{\theta}_i] \subset \mathbb{R} \forall i = 1, \dots, n$$

$$u_0(x, \theta_0) = \theta_0 y_0 + t_0$$

$$u_i(x, \theta_i) = \theta_i y_i + t_i \forall i = 1, \dots, n$$

$$f(\theta) = (y_0(\theta), \dots, y_n(\theta), t_0(\theta), \dots, t_n(\theta)).$$

2.5 Example: A combinatorial auction

This is a generalization of Example 2.3. Imagine a setting where an individual, say agent 0, is holding one unit of each of m different items ($M = \{1, 2, \dots, m\}$) and wants to sell these items. Let there be n agents who are interested in buying all the items or a non-empty subset (bundle) of the items. As before, the owner of the good is known as auctioneer and the rest of the agents are known as bidders. The auctioneer has no value for these items and just wants to sell them and this fact is a common knowledge among the bidders. Below are a few distinguishing features of this example which make it a generalization of Example 2.3.

- (i) The type set of the auctioneer is a singleton set, that is $\Theta_0 = \{\theta_0 = 0\}$. This means that auctioneer has no value for any of the bundle $A \subset M$.
- (ii) The auctioneer's type θ_0 (also known as seller's valuation for the bundles) is commonly known to all the bidders.
- (iii) The type of a bidder i consists of his valuation for every bundle of the items. Obviously, there are $2^m - 1$ possible bundles of the items, therefore, a type θ_i of a bidder i consists of $2^m - 1$ numbers.

For this example, the various building blocks of the underlying mechanism design structure are the following.

2.5a Outcome set X : An outcome in this is a vector $x = (y_i(A), t_0, t_1, \dots, t_n)_{i=1, \dots, n, A \subset M}$, where $y_i(A) = 1$ if bidder i receives the bundle $A \subset M$, $y_i(A) = 0$ otherwise. t_0 is the monetary transfer received by the auctioneer and t_i is the monetary transfer received by the bidder i . The feasibility conditions require that each bidder i must be allocated at most one bundle, that is, $\sum_{A \subset M} y_i(A) \leq 1$. Also, each object $j \in M$ must be allocated to at

most one bidder, that is, $\sum_{A \subset M | j \in A} \sum_{i=1}^n y_i(A) \leq 1$. The set of feasible alternatives is then

$$X = \left\{ (y_i(A), t_0, t_1, \dots, t_n)_{i \in N, A \subset M} \mid y_i(A) \in \{0, 1\} \forall i \in N, A \subset M; \right. \\ \left. \sum_{A \subset M} y_i(A) \leq 1 \forall i \in N; \sum_{A \subset M | j \in A} \sum_{i=1}^n y_i(A) \leq 1 \forall j \in M; \right. \\ \left. t_0 \geq 0; t_i \leq 0 \forall i \in N; \sum_{i=0}^n t_i = 0 \right\}.$$

2.5b Type set Θ_i : In this example, the type set of the auctioneer is a singleton set $\Theta_0 = \{\theta_0 = 0\}$. However, the type θ_i of each of the bidders i is a tuple $(\theta_i(A))_{A \subset M}$, where $\theta_i(A) \in [V_i, \bar{V}_i] \subset \mathbb{R} \forall A \subset M$ represents the value of agent i for bundle A . Here we are assuming that bidder i 's valuation for each bundle A lies in the interval $[V_i, \bar{V}_i]$. Therefore, the set Θ_i can be defined in the following manner:

$$\Theta_i = \{(\theta_i(A))_{A \subset M} \mid \theta_i(A) \in [V_i, \bar{V}_i] \subset \mathbb{R} \forall A \subset M\} = [V_i, \bar{V}_i]^{(2^m - 1)}.$$

2.5c Utility function $u_i(\cdot)$: The utility function of the auctioneer and bidders can be given by

$$u_0 = t_0 \\ u_i(x, \theta_i) = t_i + \sum_{A \subset M} \theta_i(A) y_i(A) \quad \forall i \in N.$$

2.5d Social choice function $f(\cdot)$: The general structure of the social choice function for this case is

$$f(\theta) = ((y_i(A, \theta))_{i \in N, A \subset M}, t_0(\theta), \dots, t_n(\theta)).$$

3. The mechanism design problem

All the above examples illustrate the concept of social choice function (SCF). However the problem of mechanism design does not end just by choosing some SCF. The social planner needs to address another problem, namely the problem of *information elicitation*. Having decided the SCF $f(\cdot)$, a trivial solution for the information elicitation problem seems to be to request the agents to reveal their types θ_i and then use them directly to compute the social outcome $x = f(\theta)$. However, given the fact that an outcome x yields a utility of $u_i(x, \theta_i)$ to agent i and agent i is a utility maximizing agent, the agents' preferences over the set X depend on the realization of the type profile $\theta = (\theta_1, \dots, \theta_n)$. For any given realization θ , different agents may prefer different outcomes. Therefore, it is not surprising if agent i reveals untruthful type, say $\hat{\theta}_i$, to the social planner because doing so may help him drive the social outcome towards a favourable choice, say \hat{x} . This phenomenon is known as the information revelation (or elicitation) problem and it is depicted in figure 2. One way in which social planner can tackle this problem is the use of an

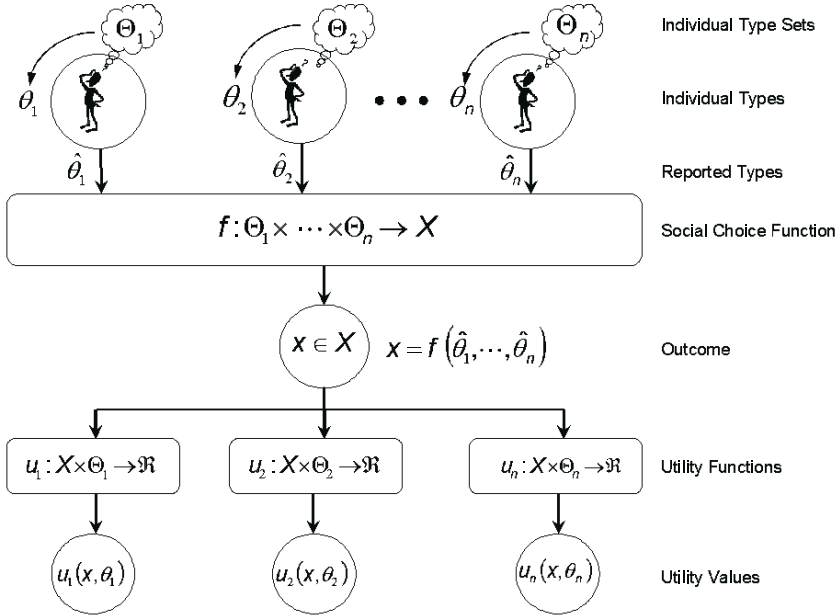


Figure 2. Information elicitation problem.

appropriate mechanism. A formal description of the mechanism design problem is provided below.

- (i) There are n individuals (or agents), indexed by $i = 1, 2, \dots, n$, who must make a collective choice from some set X , called the *outcome set*.
- (ii) Prior to the choice, however, each agent i privately observes his preferences over X . Formally, this is modelled by supposing that agent i observes a parameter, or signal θ_i that determines his preferences. The parameter θ_i is referred to as agent i 's type. The set of possible types of agent i is denoted by Θ_i .
- (iii) The agents' types, denoted by $\theta = (\theta_1, \dots, \theta_n)$ are drawn according to a probability distribution function $\Phi \in \Delta\Theta$, where $\Theta = \Theta_1 \times \dots \times \Theta_n$, and $\Delta\Theta$ is the set of all the probability distribution functions over the set Θ . Let ϕ be the corresponding probability density function.
- (iv) Each agent i is rational and intelligent and this fact is modelled by assuming that the agents always try to maximize a utility function $u_i: X \times \Theta_i \rightarrow \mathbb{R}$.
- (v) The probability density $\phi(\cdot)$, the type sets $\Theta_1, \dots, \Theta_n$, and the utility functions $u_i(\cdot)$ are assumed to be *common knowledge* among the agents. Note that the utility function $u_i(\cdot)$ of agent i depends on both the outcome x and the type θ_i . Even though, the type θ_i is not common knowledge, by saying that $u_i(\cdot)$ is common knowledge we mean that for any given type θ_i , the social planner and every other agent can evaluate the utility function of agent i .

In the above situation, the social planner faces two problems:

- (i) **Preference Aggregation Problem:** The first problem that is faced by the social planner is the following: 'For a given type profile $\theta = (\theta_1, \dots, \theta_n)$ of the agents, which outcome $x \in X$ should be chosen?'

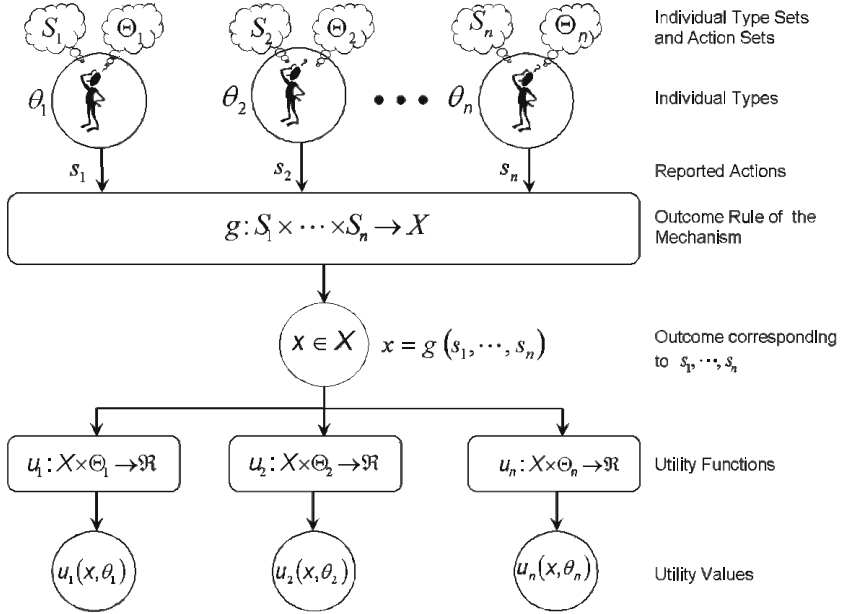


Figure 3. Abstract view of a mechanism.

- (ii) **Information Revelation (Elicitation) Problem:** Assuming that the preference aggregation problem has been solved, the other problem that is faced by the social planner is the following: ‘How do we extract the true type θ_i of each agent i , which is the private information of agent i ?’

3.1 Mechanisms

DEFINITION 3.1.

A mechanism $\mathcal{M} = ((S_i)_{i \in N}, g(\cdot))$ is a collection of action sets (S_1, \dots, S_n) and an outcome function $g: S \rightarrow X$, where $S = S_1 \times \dots \times S_n$.

The set S_i for each agent i describes the set of actions available to that agent. Based on his actual type θ_i , each agent i will choose some action, say $s_i \in S_i$. Once all the agents choose their actions, the social planner uses this profile of the actions $s = (s_1, \dots, s_n)$ to pick a social outcome $x = g(s)$. This phenomenon is depicted in figure 3. In view of the above definition, the trivial scheme of asking the agents to reveal their types becomes a special case; this special case is called a *direct revelation mechanism (DRM)*.

DEFINITION 3.2.

Given a social choice function $f: \Theta \rightarrow X$, a mechanism $\mathcal{D} = ((\Theta_i)_{i \in N}, f(\cdot))$ is known as a *direct revelation mechanism (DRM)* corresponding to $f(\cdot)$.

- Given a social choice function $f(\cdot)$, note that a direct revelation mechanism is a special case of a mechanism $\mathcal{M} = ((S_i)_{i \in N}, g(\cdot))$ with $S_i = \Theta_i \ \forall i \in N$ and $g = f$.

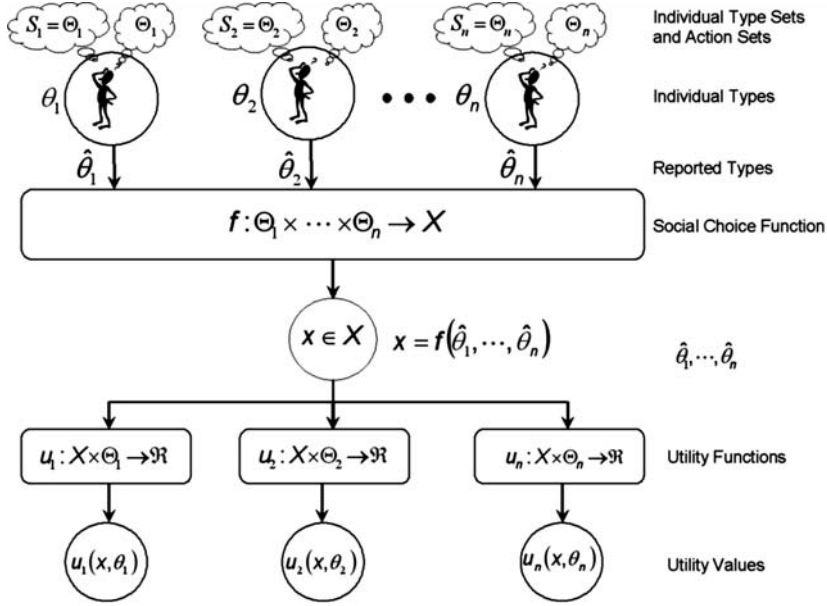


Figure 4. Abstract view of a direct revelation mechanism.

- Mechanisms that are not direct revelation mechanisms are typically referred to as *indirect mechanisms*.

Figure 4 illustrates the idea behind direct revelation mechanism. In what follows, we present a few examples of both direct revelation mechanism and indirect mechanism.

3.2 Example: Fair bilateral trade

Consider the previous bilateral trade example Example 2.2. Let us assume that the broker agent invites the seller (agent 1) to report his type θ_1 directly to him in a confidential manner. θ_1 is the minimum amount at which agent 1 is willing to sell the item. The broker agent also invites the buyer (agent 2) to report his type θ_2 directly to him in a confidential manner. θ_2 is the maximum amount that agent 2 is willing to pay for the item. Let us assume that seller reports his type to be $\hat{\theta}_1$ which may be different from actual type θ_1 and buyer reports his type to be $\hat{\theta}_2$ which also may be different from actual type θ_2 . Now the broker agent uses these reported types $\hat{\theta}_1$ and $\hat{\theta}_2$ in order to decide an outcome in the following manner. If $\hat{\theta}_1 < \hat{\theta}_2$ then the broker assigns the good to agent 2. The broker charges an amount equal to $\frac{\hat{\theta}_1 + \hat{\theta}_2}{2}$ from agent 2 and pays the same amount to the agent 1. However, if $\hat{\theta}_2 < \hat{\theta}_1$ then no trade takes place.

In this example, the mechanism that is being used by the broker agent is a direct revelation mechanism $\mathcal{D} = (\Theta_1, \Theta_2, f(\cdot))$, where the social choice function $f(\cdot)$ is given by $f(\theta) = (y_1(\theta), y_2(\theta), t_1(\theta), t_2(\theta))$, where $\theta = (\theta_1, \theta_2)$. The functions $y_i(\cdot)$ are known as *winner determination rules* and the functions $t_i(\cdot)$ are known as *payment rules*. The winner determination and payment rules are given as follows.

$$y_2(\theta) = \begin{cases} 1 & : \theta_1 < \theta_2 \\ 0 & : \text{otherwise} \end{cases}$$

$$y_1(\theta) = 1 - y_2(\theta)$$

$$t_1(\theta) = y_1(\theta) \left(\frac{\theta_1 + \theta_2}{2} \right)$$

$$t_2(\theta) = -t_1(\theta).$$

3.3 Example: First-price sealed bid auction

Consider the Example 2.3 of single unit–single item auction without reserve price. Let us assume that the auctioneer himself is a social planner and he invites each bidder i to bid an amount $b_i \geq 0$ directly to him in a confidential manner. The bid b_i means that bidder i is ready to pay an amount b_i if he receives the good. The agent i decides bid b_i based on his actual type θ_i . The bids are opened by the auctioneer and the bidder with the highest bid gets the good and pays to the auctioneer an amount equal to his bid. The other bidders pay nothing. If there are several highest bids, we suppose that the lowest numbered of these bids gets the good. We could also break the tie by randomly selecting one of the highest bidders.

In this example, the mechanism that is being used by the auctioneer is an indirect mechanism $\mathcal{M} = ((S_i)_{i \in N}, g(\cdot))$, where $S_i \subset \mathbb{R}^+$ is the set of bids that bidder i can submit to the auctioneer and $g(\cdot)$ is the outcome rule given by $g(b) = (y_1(b), \dots, y_n(b), t_1(b), \dots, t_n(b))$, where $b = (b_1, \dots, b_n)$. The functions $y_i(\cdot)$ are known as *winner determination rules* and the functions $t_i(\cdot)$ are known as *payment rules*. If we define $b^{(k)}$ to be the k^{th} highest element in (b_1, \dots, b_n) and $(b_{-i})^{(k)}$ to be the k^{th} highest element in $(b_1, \dots, b_{i-1}, b_{i+1}, \dots, b_n)$, then winner determination and payment rule can be written in following manner.

$$y_i(b) = \begin{cases} 1 & : \text{ if } b_i = b^{(1)} \\ 0 & : \text{ otherwise} \end{cases}$$

$$t_i(b) = -b_i y_i(b).$$

A few remarks are in order with regard to this example.

- While writing the above winner determination and payment rule, we have assumed that any two bidders bidding the same bid value is a zero probability event.
- If $S_i = \Theta_i = [\underline{\theta}_i, \bar{\theta}_i] \quad \forall i \in N$ then this indirect mechanism becomes a direct revelation mechanism.

3.4 Example: Second-price sealed bid (Vickrey) auction

The setting is the same as the first-price auction. The only difference here is in terms of the allocation and payment rules invoked by the auctioneer. In the second-price sealed bid auction, the bidder with the highest bid gets the good and pays to the auctioneer an amount equal to the second highest bid. The winner determination and payment rules for this auction can be given as follows.

$$y_i(b) = \begin{cases} 1 & : \text{ if } b_i = b^{(1)} \\ 0 & : \text{ otherwise} \end{cases}$$

$$t_i(b) = -(b_{-i})^{(1)} y_i(b).$$

This auction is also called as *Vickrey auction*, after the Nobel prize winning work of Vickrey (1961).

3.5 Example: Generalized Vickrey Auction (GVA)

Once again, consider the example 2.5 of single unit-multi item auction without reserve price. Let us assume that auctioneer himself is a social planner and he invites each bidder i to report bid b_i directly to him in a confidential manner. In this example, the bid structure is as follows:

$$b_i = (b_i(A))_{ACM}; b_i(A) \geq 0 \quad \forall A \subset M.$$

The bids are opened by the auctioneer and the bidders are allocated to the bundles in such a way that sum of the valuations of all the allocated bundles are maximized. Each bidder pays to the auctioneer an amount equal to his marginal contribution to the trade.

In this example, the mechanism that is being used by the auctioneer is an indirect mechanism $\mathcal{M} = ((S_i)_{i \in N}, g(\cdot))$, where $S_i \subset (\mathbb{R}^+)^{2^m - 1}$ is the set of bids that bidder i can submit to the auctioneer and $g(\cdot)$ is the outcome rule given by $g(b) = ((y_i^*(A, b))_{i \in N, ACM}, t_1(b), \dots, t_n(b))$, where $b = (b_1, \dots, b_n)$. The functions $y_i^*(\cdot, \cdot)$ are known as *winner determination rules* and the functions $t_i(\cdot)$ are known as *payment rules*. The winner determination rule $y_i^*(\cdot, \cdot)$ for this auction is basically solution of the following optimization problem.

Maximize

$$\sum_{i=1}^n \sum_{ACM} b_i(A) y_i(A, b)$$

subject to

- (i) $\sum_{ACM} y_i(A, b) \leq 1 \quad \forall i \in N$
- (ii) $\sum_{ACM | j \in A} \sum_{i=1}^n y_i(A, b) \leq 1 \quad \forall j \in M$
- (iii) $y_i(A, b) \in \{0, 1\} \quad \forall i \in N, A \subset M.$

The payment rule $t_i(\cdot)$ for this auction is given by the following relation

$$t_i(b) = \sum_{j \neq i} v_j(k^*(b), b_j) - \sum_{j \neq i} v_j(k_{-i}^*(b_{-i}), b_j),$$

where $v_j(k^*(b), b_j) = \sum_{ACM} b_j(A) y_j^*(A, b)$ is total value of the bundle which is allocated to the bidder j . The quantity $v_j(k_{-i}^*(b_{-i}), b_j) = \sum_{ACM} b_j(A) y_j^*(A, b_{-i})$ is the total value of the bundle that will be allocated to the bidder $j \neq i$ if the bidder i were not present into the system. It is easy to verify that if set M consists of just one item then above winner determination and payment rule will precisely be the same as winner determination and payment rule of the Vickrey auction, therefore, the name Generalized Vickrey Auction.

3.6 Bayesian game induced by a mechanism

In view of the definition of the indirect mechanism and direct revelation mechanism, we can say that a social planner can either use an indirect mechanism \mathcal{M} or a direct mechanism \mathcal{D} to elicit the information about the agents' preferences in an indirect or a direct manner, respectively. As we assumed earlier, all the agents are rational and intelligent. Therefore, after knowing about the mechanism $\mathcal{M} = ((S_i)_{i \in N}, g(\cdot))$ chosen by the social planner, each agent i starts doing an analysis regarding which action s_i will result in his most favourable outcome

and comes up with a strategy $s_i: \Theta_i \rightarrow S_i$ to take the action. This phenomenon leads to a game among the agents. A mechanism $\mathcal{M} = ((S_i)_{i \in N}, g(\cdot))$ combined with possible types of the agents $(\Theta_1, \dots, \Theta_n)$, probability density $\phi(\cdot)$, and utility functions $(u_1(\cdot), \dots, u_n(\cdot))$ defines a Bayesian game of incomplete information that gets induced among the agents when the social planner invokes this mechanism as a means to solve the information elicitation problem. The induced Bayesian game Γ^b is given in the following manner:

$$\Gamma^b = (N, (S_i)_{i \in N}, (\Theta_i)_{i \in N}, \phi(\cdot), (\bar{u}_i)_{i \in N}),$$

where $\bar{u}_i: S \times \Theta \rightarrow \mathbb{R}$ is the utility function of agent i and is defined in following manner

$$\bar{u}_i(s, \theta) = u_i(g(s), \theta_i),$$

where $S = \prod_{i \in N} S_i$, and $\Theta = \prod_{i \in N} \Theta_i$.

Knowing the fact that choosing any mechanism $\mathcal{M} = ((S_i)_{i \in N}, g(\cdot))$ will induce the game among the agents, and the agents will respond to it in a way suggested by the corresponding equilibrium strategy of the game, the social planner now worries about whether or not the outcome of the game matches with the outcome of the social choice function $f(\theta)$ (if all the agents had revealed their true types when asked directly). This notion is captured in the definition that follows.

4. Implementing a social choice function

DEFINITION 4.1.

We say that the mechanism $\mathcal{M} = ((S_i)_{i \in N}, g(\cdot))$ implements the social choice function $f(\cdot)$ if there is a pure strategy equilibrium $s^*(\cdot) = (s_1^*(\cdot), \dots, s_n^*(\cdot))$ of the Bayesian game Γ^b induced by \mathcal{M} such that $g(s_1^*(\theta_1), \dots, s_n^*(\theta_n)) = f(\theta_1, \dots, \theta_n) \forall (\theta_1, \dots, \theta_n) \in \Theta$.

The figure 5 explains the idea behind what we mean by mechanism implementing a social choice function. Depending on the underlying equilibrium concept, two ways of implementing an SCF $f(\cdot)$ are standard in the literature.

4.1 Implementing a social choice function in dominant strategy equilibrium

First, we define the notion of a weakly dominant strategy equilibrium of the Bayesian game Γ^b .

DEFINITION 4.2 (Weakly dominant strategy equilibrium).

A pure strategy profile $s^d(\cdot) = (s_1^d(\cdot), \dots, s_n^d(\cdot))$ of the game Γ^b induced by the mechanism \mathcal{M} , is said to be a weakly dominant strategy equilibrium if it satisfies the following condition.

$$u_i(g(s_i^d(\theta_i), s_{-i}(\theta_{-i})), \theta_i) \geq u_i(g(s_i'(\theta_i), s_{-i}(\theta_{-i})), \theta_i) \\ \forall i \in N, \forall \theta_i \in \Theta_i, \forall \theta_{-i} \in \Theta_{-i}, \forall s_i'(\cdot) \in S_i, \forall s_{-i}(\cdot) \in S_{-i}, \quad (1)$$

where S_i is the set of pure strategies of the agent i in the induced Bayesian game Γ^b , and S_{-i} is the set of pure strategy profiles of all the agents except agent i .

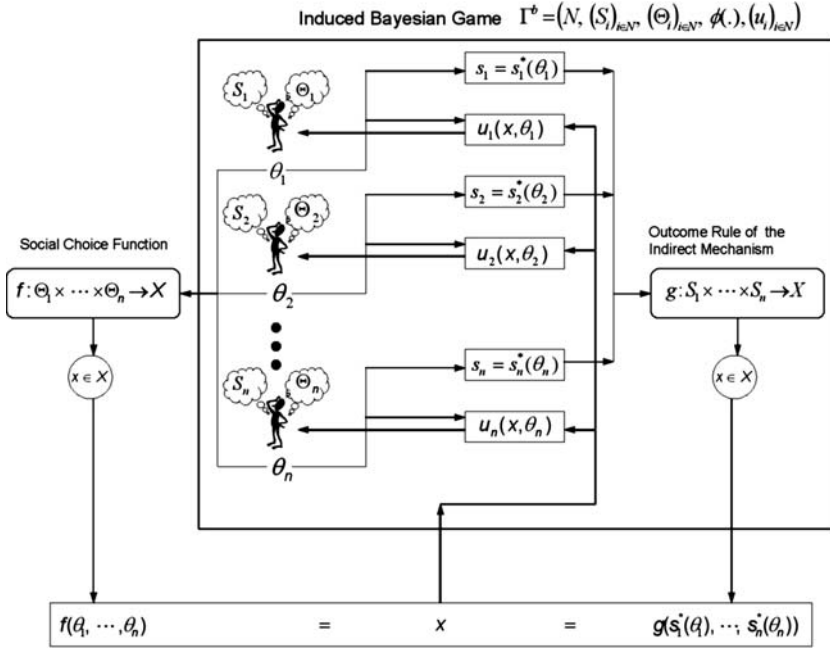


Figure 5. Mechanism $\mathcal{M} = ((S_i)_{i \in N}, g(\cdot))$ implements the social choice function $f(\cdot)$.

DEFINITION 4.3. We say that the mechanism $\mathcal{M} = ((S_i)_{i \in N}, g(\cdot))$ implements the social choice function $f(\cdot)$ in dominant strategy equilibrium if there is a weakly dominant strategy equilibrium $s^d(\cdot) = (s_1^d(\cdot), \dots, s_n^d(\cdot))$ of the game Γ^b induced by \mathcal{M} such that

$$g(s_1^d(\theta_1), \dots, s_n^d(\theta_n)) = f(\theta_1, \dots, \theta_n) \quad \forall (\theta_1, \dots, \theta_n) \in \Theta.$$

4.2 Implementing a social choice function in Bayesian–Nash equilibrium

DEFINITION 4.4.

We say that the mechanism $\mathcal{M} = ((S_i)_{i \in N}, g(\cdot))$ implements the social choice function $f(\cdot)$ in Bayesian–Nash equilibrium if there is a pure strategy Bayesian–Nash equilibrium $s^*(\cdot) = (s_1^*(\cdot), \dots, s_n^*(\cdot))$ of the game Γ^b induced by \mathcal{M} such that

$$g(s_1^*(\theta_1), \dots, s_n^*(\theta_n)) = f(\theta_1, \dots, \theta_n) \quad \forall (\theta_1, \dots, \theta_n) \in \Theta.$$

Following is the definition of Bayesian–Nash equilibrium of the Bayesian game Γ^b .

DEFINITION 4.5 (Bayesian–Nash equilibrium).

A pure strategy profile $s^*(\cdot) = (s_1^*(\cdot), \dots, s_n^*(\cdot))$ of the game Γ^b induced by the mechanism \mathcal{M} , is a Bayesian–Nash equilibrium if it satisfies the following condition.

$$E_{\theta_{-i}}[u_i(g(s_i^*(\theta_i), s_{-i}^*(\theta_{-i})), \theta_i) | \theta_i] \geq E_{\theta_{-i}}[u_i(g(s_i'(\theta_i), s_{-i}^*(\theta_{-i})), \theta_i) | \theta_i]$$

$$\forall i \in N, \forall \theta_i \in \Theta_i, \forall s_i'(\cdot) \in S_i. \quad (2)$$

Following is a proposition that establishes the relationship between two equilibrium concepts defined above. The proof is straightforward.

PROPOSITION 4.1.

A weakly dominant strategy equilibrium $s^d(\cdot) = (s_1^d(\cdot), \dots, s_n^d(\cdot))$ of the Bayesian game Γ^b induced by \mathcal{M} , is always a pure strategy Bayesian–Nash equilibrium of the same Bayesian game Γ^b .

COROLLARY 4.1.

If the mechanism $\mathcal{M} = ((S_i)_{i \in N}, g(\cdot))$ implements the social choice function $f(\cdot)$ in dominant strategy equilibrium, then it also implements $f(\cdot)$ in Bayesian–Nash equilibrium.

In what follows, we offer a few caveats to these definitions of implementing a social choice function.

- (i) The game Γ^b induced by the mechanism \mathcal{M} may have more than one equilibrium, but the above definition requires only that *one of them* induces outcomes in accordance with the SCF $f(\cdot)$. Implicitly, then, the above definition assumes that, if multiple equilibria exist, the agents will play the equilibrium that the mechanism designer (social planner) wants.
- (ii) Another implicit assumption of the above definition is that the game induced by the mechanism is a simultaneous move game, that is all the agents, after learning their types, choose their actions simultaneously. However, it is quite possible that the mechanism forces some agent(s) to lead the game, taking the action first followed by the actions of the remaining agents. In such a case, the game induced by the mechanism becomes a *Stackelberg game* (Garg & Narahari 2005, Baar & Olsder 1999).

5. Properties of a social choice function

We have seen that a mechanism provides a solution to both the problem of information elicitation and the problem of preferences aggregation if it can implement the desired social choice function $f(\cdot)$. It is obvious that some SCFs are implementable and some are not. Before we look into the question of characterizing the space of implementable social choice functions, it is important to know which social choice function ideally a social planner would prefer to be implemented. In this section, we highlight a few properties of an SCF which ideally a social planner would wish the SCF to have.

Note that the fundamental characteristic of a social planner is that he is neutral to all the agents. Therefore, it is obvious for the social planner to be concerned about whether the outcome $f(\theta_1, \dots, \theta_n)$ is socially fair or not. For this, a social planner would always like to use an SCF $f(\cdot)$ which satisfies as many desirable properties from the perspective of fairness as possible. A few important properties, which ideally a social planner would want an SCF $f(\cdot)$ to satisfy, are the following.

5.1 Ex-post efficiency

DEFINITION 5.1.

The SCF $f: \Theta \rightarrow X$ is said to be ex-post efficient (or Paretian) if for any profile of agents' types $\theta = (\theta_1, \dots, \theta_n)$, and any pair of alternatives $x, y \in X$, such that $u_i(x, \theta_i) \geq u_i(y, \theta_i) \forall i$ and $u_i(x, \theta_i) > u_i(y, \theta_i)$ for some i , we have $y \neq f(\theta_1, \dots, \theta_n)$.

An alternative definition of ex-post efficiency can be given in the following manner.

DEFINITION 5.2.

The SCF $f: \Theta \rightarrow X$ is said to be ex-post efficient if for no profile of agents' types $\theta = (\theta_1, \dots, \theta_n)$ does there exist an $x \in X$ such that $u_i(x, \theta_i) \geq u_i(f(\theta), \theta_i) \forall i$ and $u_i(x, \theta_i) > u_i(f(\theta), \theta_i)$ for some i .

5.2 Non-dictatorial SCF

We define this through a dictatorial SCF.

DEFINITION 5.3.

The SCF $f: \Theta \rightarrow X$ is said to be dictatorial if for every profile of agents' type $\theta = (\theta_1, \dots, \theta_n)$, we have $f(\theta_1, \dots, \theta_n) \in \{x \in X | u_d(x, \theta_d) \geq u_d(y, \theta_d) \forall y \in X\}$, where d is a particular agent known as dictator.

An SCF $f: \Theta \rightarrow X$ is said to be non-dictatorial if it is not dictatorial.

5.3 Incentive compatibility (IC)

DEFINITION 5.4.

The SCF $f(\cdot)$ is said to be incentive compatible (or truthfully implementable) if the direct revelation mechanism $\mathcal{D} = ((\Theta_i)_{i \in N}, f(\cdot))$ has a pure strategy equilibrium $s^*(\cdot) = (s_1^*(\cdot), \dots, s_n^*(\cdot))$ in which $s_i^*(\theta_i) = \theta_i, \forall \theta_i \in \Theta_i, \forall i \in N$.

That is, truth telling by each agent constitutes an equilibrium of the game induced by \mathcal{D} . It is easy to verify that if an SCF $f(\cdot)$ is incentive compatible then the direct revelation mechanism $\mathcal{D} = ((\Theta_i)_{i \in N}, f(\cdot))$ can implement it. That is, directly asking the agents to report their types and plugging this information in $f(\cdot)$ to get the social outcome will solve both the problem of information elicitation and the problem of preferences aggregation.

The notion of incentive compatibility, which is central to mechanism design theory, was first introduced by Hurwicz (1972). Based on the type of equilibrium concept used, two types of incentive compatibility are given below.

5.4 Dominant strategy incentive compatibility (DSIC)

DEFINITION 5.5.

The SCF $f(\cdot)$ is said to be dominant strategy incentive compatible (or truthfully implementable in dominant strategies)¹ if the direct revelation mechanism $\mathcal{D} = ((\Theta_i)_{i \in N}, f(\cdot))$ has a *dominant strategy equilibrium* $s^d(\cdot) = (s_1^d(\cdot), \dots, s_n^d(\cdot))$ in which $s_i^d(\theta_i) = \theta_i, \forall \theta_i \in \Theta_i, \forall i \in N$.

That is, truth telling by each agent constitutes a dominant strategy equilibrium of the game induced by \mathcal{D} . Following is a necessary and sufficient condition for an SCF $f(\cdot)$ to be dominant strategy incentive compatible:

$$u_i(f(\theta_i, \theta_{-i}), \theta_i) \geq u_i(f(\hat{\theta}_i, \theta_{-i}), \theta_i), \forall i \in N, \forall \theta_i \in \Theta_i, \forall \theta_{-i} \in \Theta_{-i}, \forall \hat{\theta}_i \in \Theta_i. \quad (3)$$

The above condition says that if the SCF $f(\cdot)$ is DSIC, then, irrespective of what the other agents are doing, it is always in the best interest of agent i to report his true type θ_i .

¹Strategy-proof, cheat-proof, straightforward are the alternative phrases used for this property.

5.5 Bayesian incentive compatibility (BIC)

DEFINITION 5.6.

The SCF $f(\cdot)$ is said to be Bayesian incentive compatible (or truthfully implementable in Bayesian–Nash equilibrium) if the direct revelation mechanism $\mathcal{D} = ((\Theta_i)_{i \in N}, f(\cdot))$ has a *Bayesian–Nash equilibrium* $s^*(\cdot) = (s_1^*(\cdot), \dots, s_n^*(\cdot))$ in which $s_i^*(\theta_i) = \theta_i, \forall \theta_i \in \Theta_i, \forall i \in N$.

That is, truth telling by each agent constitutes a Bayesian–Nash equilibrium of the game induced by \mathcal{D} . Following is a necessary and sufficient condition for an SCF $f(\cdot)$ to be Bayesian incentive compatible:

$$\begin{aligned} E_{\theta_{-i}}[u_i(f(\theta_i, \theta_{-i}), \theta_i) | \theta_i] \\ \geq E_{\theta_{-i}}[u_i(f(\hat{\theta}_i, \theta_{-i}), \theta_i) | \theta_i], \forall i \in N, \forall \theta_i \in \Theta_i, \forall \hat{\theta}_i \in \Theta_i. \end{aligned} \quad (4)$$

The following proposition illustrates the relationship between these two notions of incentive compatibility of a social choice function. The proof of this proposition is quite straightforward.

PROPOSITION 5.1.

If a social choice function $f(\cdot)$ is dominant strategy incentive compatible then it is also Bayesian incentive compatible.

6. The revelation principle

This is one of the most fundamental results in the theory of mechanism design. This principle basically illustrates the relationship between an indirect mechanism \mathcal{M} and a direct revelation mechanism \mathcal{D} for any SCF $f(\cdot)$. This result enables us to restrict our inquiry about truthful implementation of an SCF to the class of direct revelation mechanisms only.

6.1 The revelation principle for dominant strategy equilibrium

PROPOSITION 6.1.

Suppose that there exists a mechanism $\mathcal{M} = (S_1, \dots, S_n, g(\cdot))$ that implements the social choice function $f(\cdot)$ in dominant strategy equilibrium. Then $f(\cdot)$ is truthfully implementable in dominant strategy equilibrium (dominant strategy incentive compatible).

Proof. If $\mathcal{M} = (S_1, \dots, S_n, g(\cdot))$ implements $f(\cdot)$ in dominant strategy, then there exists a profile of strategies $s^d(\cdot) = (s_1^d(\cdot), \dots, s_n^d(\cdot))$ such that

$$g(s_1^d(\theta_1), \dots, s_n^d(\theta_n)) = f(\theta_1, \dots, \theta_n) \quad \forall (\theta_1, \dots, \theta_n) \in \Theta \quad (5)$$

and

$$\begin{aligned} u_i(g(s_i^d(\theta_i), s_{-i}(\theta_{-i})), \theta_i) \geq u_i(g(s_i'(\theta_i), s_{-i}(\theta_{-i})), \theta_i) \\ \forall i \in N, \forall \theta_i \in \Theta_i, \forall \theta_{-i} \in \Theta_{-i}, \forall s_i'(\cdot) \in S_i, \forall s_{-i}(\cdot) \in S_{-i}. \end{aligned} \quad (6)$$

Condition (6) implies, in particular, that

$$u_i(g(s_i^d(\theta_i), s_{-i}^d(\theta_{-i})), \theta_i) \geq u_i(g(s_i^d(\hat{\theta}_i), s_{-i}^d(\theta_{-i})), \theta_i) \quad \forall i \in N, \forall \theta_i \in \Theta_i, \forall \hat{\theta}_i \in \Theta_i, \forall \theta_{-i} \in \Theta_{-i}. \quad (7)$$

Conditions (5) and (7) together implies that

$$u_i(f(\theta_i, \theta_{-i}), \theta_i) \geq u_i(f(\hat{\theta}_i, \theta_{-i}), \theta_i), \forall i \in N, \forall \theta_i \in \Theta_i, \forall \theta_{-i} \in \Theta_{-i}, \forall \hat{\theta}_i \in \Theta_i.$$

But this is precisely condition (3), the condition for $f(\cdot)$ to be truthfully implementable in dominant strategies. Q.E.D.

The idea behind the revelation principle can be understood with the help of figure 6. The set I_{DSE} consists of all the social choice functions that are implementable by some indirect mechanism \mathcal{M} in dominant strategy equilibrium. The set $DSIC$ consists of all the social choice functions that are truthfully implementable by some direct mechanism \mathcal{D} in dominant strategy equilibrium. Recall that a direct mechanism \mathcal{D} can also be viewed as an indirect mechanism. Therefore, it is obvious that we must have

$$DSIC \subset I_{DSE}. \quad (8)$$

In view of relation 8, we can say that the revelation principle for dominant strategy equilibrium basically says that $I_{DSE} \subset DSIC$, which further implies that $I_{DSE} = DSIC$.

Thus, the revelation principle for dominant strategy equilibrium says that an SCF $f(\cdot)$ is implementable by an indirect mechanism $\mathcal{M} = (S_1, \dots, S_n, g(\cdot))$ in dominant strategy equilibrium if it is truthfully implementable by the mechanism $\mathcal{D} = ((\Theta_i)_{i \in N}, f(\cdot))$ in dominant strategy equilibrium.

6.2 The revelation principle for Bayesian–Nash equilibrium

PROPOSITION 6.2.

Suppose that there exists a mechanism $\mathcal{M} = (S_1, \dots, S_n, g(\cdot))$ that implements the social choice function $f(\cdot)$ in Bayesian–Nash equilibrium. Then $f(\cdot)$ is truthfully implementable in Bayesian–Nash equilibrium (Bayesian incentive compatible).

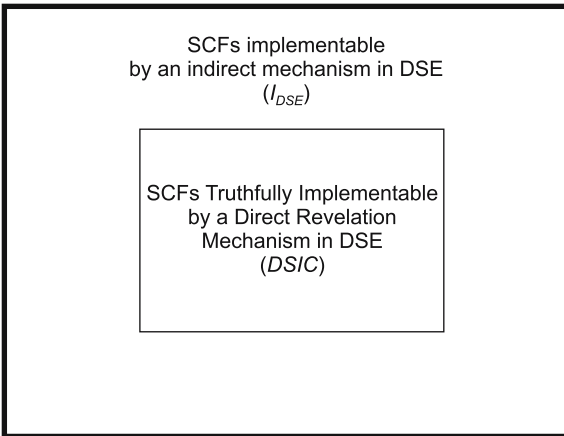


Figure 6. Revelation principle for dominant strategy equilibrium.

Proof. If $\mathcal{M} = (S_1, \dots, S_n, g(\cdot))$ implements $f(\cdot)$ in Bayesian–Nash equilibrium, then there exists a profile of strategies $s^*(\cdot) = (s_1^*(\cdot), \dots, s_n^*(\cdot))$ such that

$$g(s_1^*(\theta_1), \dots, s_n^*(\theta_n)) = f(\theta_1, \dots, \theta_n) \quad \forall (\theta_1, \dots, \theta_n) \in \Theta \quad (9)$$

and

$$E_{\theta_{-i}} [u_i(g(s_i^*(\theta_i), s_{-i}^*(\theta_{-i})), \theta_i) | \theta_i] \geq E_{\theta_{-i}} [u_i(g(s_i'(\theta_i), s_{-i}^*(\theta_{-i})), \theta_i) | \theta_i] \\ \forall i \in N, \forall \theta_i \in \Theta_i, \forall s_i'(\cdot) \in S_i. \quad (10)$$

Condition (10) implies, in particular, that

$$E_{\theta_{-i}} [u_i(g(s_i^*(\theta_i), s_{-i}^*(\theta_{-i})), \theta_i) | \theta_i] \geq E_{\theta_{-i}} [u_i(g(s_i^*(\hat{\theta}_i), s_{-i}^*(\theta_{-i})), \theta_i) | \theta_i] \\ \forall i \in N, \forall \theta_i \in \Theta_i, \forall \hat{\theta}_i \in \Theta_i. \quad (11)$$

Conditions (9) and (11) together implies that

$$E_{\theta_{-i}} [u_i(f(\theta_i, \theta_{-i}), \theta_i) | \theta_i] \\ \geq E_{\theta_{-i}} [u_i(f(\hat{\theta}_i, \theta_{-i}), \theta_i) | \theta_i], \forall i \in N, \forall \theta_i \in \Theta_i, \forall \hat{\theta}_i \in \Theta_i.$$

But this is precisely condition (4), the condition for $f(\cdot)$ to be truthfully implementable in Bayesian–Nash equilibrium. Q.E.D.

In a way similar to the revelation principle for dominant strategy equilibrium, the revelation principle for Bayesian–Nash equilibrium can be explained with the help of figure 7. Following a similar line of arguments, we can get the following relation for this case also:

$$BIC \subset I_{BNE}. \quad (12)$$

In view of relation (12), we can say that the revelation principle for Bayesian–Nash equilibrium basically says that $I_{BNE} \subset BIC$, which further implies that $I_{BNE} = BIC$.

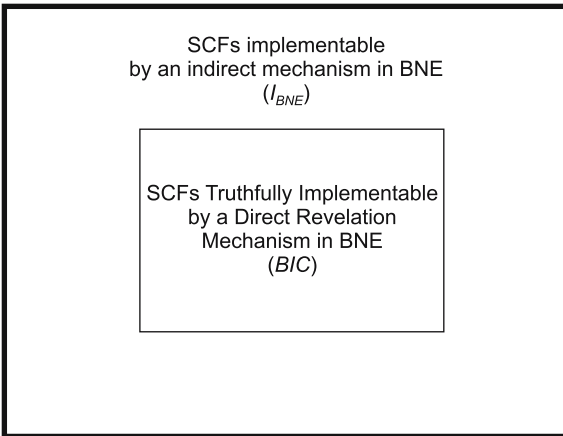


Figure 7. Revelation principle for Bayesian–Nash equilibrium.

Thus, the revelation principle for Bayesian–Nash equilibrium says that an SCF $f(\cdot)$ is implementable by an indirect mechanism $\mathcal{M} = (S_1, \dots, S_n, g(\cdot))$ in Bayesian–Nash equilibrium if it is truthfully implementable by the mechanism $\mathcal{D} = ((\Theta_i)_{i \in N}, f(\cdot))$ in Bayesian–Nash equilibrium.

In view of the above revelation principle, from now onwards, we will be just focusing on direct revelation mechanisms without loss of any generality.

7. The Gibbard–Satterthwaite impossibility theorem

Ideally, a social planner would prefer to implement a social choice function $f(\cdot)$ which is *ex-post efficient*, *non-dictatorial*, and *dominant strategy incentive compatible*. Now the question is: Does there exist any such social choice function? The answer is *no*.

The Gibbard–Satterthwaite impossibility theorem shows that for a very general class of problems there is no hope of implementing any satisfactory social choice function in dominant strategies. This is an important landmark result which has shaped the course of research on incentives and implementation to a great extent and it was discovered independently by Gibbard (1973) and Satterthwaite (1975). To understand the precise statement of this theorem, we need to build up a few concepts.

7.1 Utility function and preference relation

We have already seen that a given preference of an agent i , over the outcome set X , can also be described by means of a *utility function* $u_i: X \rightarrow \mathbb{R}$ which assigns a numerical value to each element in X . A utility function u_i always induces a *unique* preference relation \succsim_i on X which can be described in following manner

$$x \succsim_i y \Leftrightarrow u_i(x) \geq u_i(y).$$

The following proposition establishes the relationship between these two ways of expressing the preferences of an agent i over the set X - *preference relation and utility function*. The above preference relation is often called a *rational preference relation* and it is formally defined as follows.

DEFINITION 7.1 (Rational preference relation).

We say that a relation \succsim_i on the set X is called a *rational preference relation* if it possesses the following three properties:

- (i) Reflexivity: $\forall x \in X$, we have $x \succsim_i x$.
- (ii) Completeness: $\forall x, y \in X$, we have that $x \succsim_i y$ or $y \succsim_i x$ (or both).
- (iii) Transitivity: $\forall x, y, z \in X$, if $x \succsim_i y$ and $y \succsim_i z$, then $x \succsim_i z$.

PROPOSITION 7.1.

- (i) If a preference relation \succsim_i over X can be induced by some utility function $u_i(\cdot)$, then it will be a rational preference relation.
- (ii) For every preference relation $\succsim_i \in \mathcal{R}$, there need not exist a utility function which will induce it. However, it is true when the set X is finite.

- (iii) For a given preference relation $\succsim_i \in \mathcal{R}$, if there exists a utility function which induces it, then this need not be the unique one. Indeed, if the utility function $u_i(\cdot)$ induces \succsim_i , then $v(x) = f(u_i(x))$ is another utility function which will also induce \succsim_i , where $f: \mathbb{R} \rightarrow \mathbb{R}$ is a strictly increasing function.

DEFINITION 7.2 (Strict-total preference relation).

We say that a rational preference relation \succsim_i is strict-total if it possesses the antisymmetry property, in addition to the reflexivity, completeness, and transitivity. By antisymmetry, we mean that, for any $x, y \in X$ such that $x \neq y$, we have either $x \succsim_i y$ or $y \succsim_i x$, but not both.

The set of all the rational preference relations and strict-total preference relations over the set X are denoted by \mathcal{R} and \mathcal{S} , respectively. It is easy to see that $\mathcal{S} \subset \mathcal{R}$. The strict-total order relation is also known as *linear order relation* because it satisfies the properties of the usual ‘greater than or equal to’ relationship on the real line.

7.2 Set of ordinal preference relations

In the mechanism design problem, for agent i , the preference over the set X is described in the form of a utility function $u_i: X \times \Theta_i \rightarrow \mathbb{R}$. That is, for every possible type $\theta_i \in \Theta_i$ of agent i , we can define a utility function $u_i(\cdot, \theta_i)$ over the set X . Let this utility function induce a rational preference relation $\succsim_i(\theta_i)$ over X . The set $\mathcal{R}_i = \{\succsim_i : \succsim_i = \succsim_i(\theta_i) \text{ for some } \theta_i \in \Theta_i\}$ is known as the set of ordinal preference relations for agent i . It is easy to see that $\mathcal{R}_i \subset \mathcal{R} \forall i = 1, \dots, n$.

Now we can state the Gibbard–Satterthwaite impossibility theorem.

Theorem 7.1 (Gibbard–Satterthwaite Impossibility Theorem). *Suppose that*

- (i) *The outcome set X is finite and contains at least three elements*
- (ii) $\mathcal{R}_i = \mathcal{S} \forall i = 1, \dots, n$
- (iii) $f(\Theta) = X$, that is, the image of SCF $f(\cdot)$ is the set X .

Then the social choice function $f(\cdot)$ is truthfully implementable in dominant strategies if it is dictatorial.

Here is a brief outline of the proof of the theorem. For a detailed proof, refer to Proposition 23.C.3 of Mas-Colell *et al* (1995). To prove the necessity, we assume that the social choice function $f(\cdot)$ is dictatorial and it is shown that $f(\cdot)$ is DSIC. This is fairly straightforward. The proof of the sufficiency part of the theorem proceeds in three stages. We start with the assumption that $f(\cdot)$ is DSIC and in stage 1, we prove that $f(\cdot)$ satisfies a monotonicity property. Next, in stage 2, it is shown that $f(\cdot)$ is ex-post efficient. Finally, in stage 3, it is shown that monotonicity and ex-post efficiency together imply that $f(\cdot)$ is dictatorial.

The Gibbard–Satterthwaite impossibility theorem gives a disappointing piece of news and the question facing a social planner is what kind of SCF to look for in the face of this impossibility result. There are two possible routes that one may take:

- (i) The first route is to focus on some restricted environment where at least one of the three requirements of the Gibbard–Satterthwaite impossibility theorem is not fulfilled. *Quasi-linear environment* is one such environment where the second condition of this theorem is not satisfied and in fact all the social choice functions in such environments are non-dictatorial.

- (ii) The other route is to weaken the implementation concept and look for an SCF which is *ex-post efficient, non-dictatorial, and Bayesian incentive compatible*.

The Gibbard–Satterthwaite impossibility theorem happens to be a special case of the celebrated Arrow’s impossibility theorem in welfare economics. We now provide a discussion of this famous result in the next section before returning to discuss the above two routes for circumventing the difficulties arising out of the Gibbard–Satterthwaite theorem and the Arrow’s impossibility theorem.

8. Arrow’s impossibility theorem

Suppose a person has to select one of three fruits, A, B, or C to purchase. The agent prefers to have A over B, prefers B to C. The agent can now decide which fruit he would purchase, consistent with his preferences. Suppose there are two friends who wish to purchase one fruit out of these three. They may have different preferences, so they have to take a collective decision which is suitable for both of them. The decision is not straightforward anymore. A social welfare function is useful for describing the collective (or social) preference in such scenarios.

A social planner seeks to implement a desirable system wide solution. To have a socially desirable outcome, the social planner has to aggregate the preferences of every individual, over the outcome set X . A function, which the social planner uses to aggregate individual preference into a social preference is referred to as *social welfare function*, formally.

DEFINITION 8.1 (Social welfare function).

A social welfare function is a function $W: \mathcal{R}^n \rightarrow \mathcal{R}$ that aggregates the preference profile $(\succsim_1, \succsim_2, \dots, \succsim_n) \in \mathcal{R}^n$ of individual preferences into a public or social preference $\succsim \in \mathcal{R}$.

Note the difference between a social welfare function and social choice function. Social welfare function aggregates the individual preferences into public preference where as SCF aggregates the public preferences into public choice, that is, selects outcome from X . The Arrow’s theorem is a celebrated result in welfare economics which essentially says that a social welfare function that is required to satisfy two desirable properties, namely *unanimity* and *independence of irrelevant alternatives* (IIA) will necessarily belong to a restricted class of functions called *dictatorial* social welfare functions. We first define unanimity (also called Paretian), IIA, and dictatorship.

DEFINITION 8.2 (Unanimity).

A social welfare function W is said to satisfy unanimity (or Paretian) if for every $\succsim \in \mathcal{S}$, $W(\succsim, \dots, \succsim) = \succsim$.

Unanimity means that if all agents have identical preferences then the social preference is the same as the preference of each agent. As a simple example, in the case of the fruits situation described above, if both the agents prefer A to B and B to C, then the social preference also will prefer A to B and B to C.

DEFINITION 8.3 (Independence of irrelevant alternatives).

A social welfare function $W: \mathcal{A} \rightarrow \mathcal{R}$ defined on the domain $\mathcal{A} \subseteq \mathcal{R}^n$, is said to satisfy independence of irrelevant alternatives conditions if the social preference between any two alternatives depends only on the profile of individual preferences over the same alternatives. Formally, for any pair, $x, y \in X$ and for any pair of preference profiles $(\succsim_1, \succsim_2, \dots, \succsim_n), (\succsim'_1, \succsim'_2, \dots, \succsim'_n)$, if we denote $\succsim = W(\succsim_1, \succsim_2, \dots, \succsim_n)$ and $\succsim' = W(\succsim'_1, \succsim'_2, \dots, \succsim'_n)$, then

$$x \succsim_i y \Leftrightarrow x \succsim'_i y \quad \forall i \in N \text{ implies } x \succsim y \Leftrightarrow x \succsim' y.$$

This condition is quite a natural one. Independence of irrelevant alternatives means that the social preference of two alternatives x, y should depend upon individual preferences over x, y and not on other outcomes. This condition captures a certain consistency property, lack of which may lead to strategic manipulation. The characteristic of independence of irrelevant alternatives is also called *pairwise independence*.

DEFINITION 8.4 (Dictatorship).

A social welfare function W is said to be dictatorial or is called a dictatorship if there exists a dictator. A player $i \in N$ is called a dictator if

$$W(\succsim_1, \dots, \succsim_n) = \succsim_i \quad \forall \succsim_1, \dots, \succsim_n \in \mathcal{R}.$$

Theorem 8.1 (Arrow's impossibility theorem). *Suppose $|X| \geq 3$ and that the domain of admissible individual profiles denoted by \mathcal{A} is either \mathcal{R}^n or \mathcal{S}^n . Then every social welfare function $W: \mathcal{A} \rightarrow \mathcal{R}$ that is unanimous and satisfies independence of irrelevant alternatives, is dictatorial.*

We make a quick observation that the above theorem is a negative result for a social planner. The implication of the theorem is quite devastating: there is no social welfare function in unrestricted domain which simultaneously satisfies independence of irrelevant alternatives, unanimity, and non-dictatorial properties.

There are many proofs of this theorem. Interested readers can refer to the Section 21·C/22·D of the book by Mas-Colell *et al* (1995). An elegant proof is available in Nisan (2007). The proof proceeds in two parts. Let W be a social welfare function that satisfies unanimity and independence of irrelevant alternatives. Part 1 shows *pairwise neutrality*, by which we mean that the same social ranking rule is taken within any pair of alternatives. Part 2 invokes pairwise neutrality to show that there exists an agent i_d who is a dictator.

8.1 Example: The condorcet paradox

Suppose there are three agents who wish to buy one fruit out of three fruits, A, B, C , so, $X = \{A, B, C\}$. Suppose that to decide social preference between any two alternatives, we implement *majority voting* between those two alternatives. That is, if 2 out of 3 prefer to go for fruit A over fruit B , we will put A to be preferable to B as the social decision. It can be verified that the majority voting function satisfies unanimity and IIA. Assume that the preferences of the three agents are: (1) $B \succsim_1 A \succsim_1 C$; (2) $A \succsim_2 C \succsim_2 B$; (3) $C \succsim_3 B \succsim_3 A$. Let $\succsim = W(\succsim_1, \succsim_2, \succsim_3)$, where W is the majority voting function. Now note that $B \succsim A$ as B is preferred over A by 2 out of 3 agents. Similarly, $A \succsim C$, $C \succsim B$.

But then we cannot decide any social ordering among these 3 alternatives at this preference profile since that would imply $B \succsim A \succsim C \succsim B$, which is contradiction.²

We could not implement this social welfare function because majority voting is not a dictatorship. This exemplifies the Arrow's theorem which says that a social welfare function that satisfies unanimity and IIA is necessarily dictatorial.

Now, consider another social welfare function Z which is such that, if for any pair $(a, b) \in X$ such that $a \succeq_1 b$, Z prefers a over b . Clearly, this social welfare function satisfies unanimity and IIA and this can be implemented. Here, agent 1 is a dictator and hence the function Z is a dictatorial social welfare function.

8.2 Arrow's theorem and the Gibbard–Satterthwaite theorem

The Gibbard–Satterthwaite theorem asserts the dictatorship of social choice functions while Arrow's theorem asserts the dictatorship of social welfare functions. The Gibbard–Satterthwaite theorem is a special case of the Arrow's theorem, as shown by the following argument. Start with a non-dictatorial SCF such that $f(\Theta) = X$, that is, the image of SCF $f(\cdot)$ is the set X and $|X|$ is at least 3. One can construct a social welfare function from this social choice function and show that it is non-dictatorial, which immediately contradicts Arrow's theorem. A detailed proof of the Gibbard–Satterthwaite theorem as a special case of Arrow's theorem can be found in Nisan (2007).

9. Quasi-linear environments

This is a special and much studied class of environments where the Gibbard–Satterthwaite theorem does not hold good. In this environment, an alternative $x \in X$ is a vector of the form $x = (k, t_1, \dots, t_n)$, where k is an element of a set K , to be called as 'project choice'. The set K is a compact subset of a topological space. $t_i \in \mathbb{R}$ is a monetary transfer to agent i . If $t_i > 0$ then agent i will receive the money and if $t_i < 0$ then agent i will pay the money. We assume that we are dealing with a closed system in which n agents have no outside source of funding, i.e. $\sum_{i=1}^n t_i \leq 0$. This condition is known as *weak budget balance* condition. The set of alternatives X is therefore

$$X = \left\{ (k, t_1, \dots, t_n) : k \in K, t_i \in \mathbb{R} \forall i = 1, \dots, n, \sum_i t_i \leq 0 \right\}.$$

A social choice function in this quasi-linear environment takes the form $f(\theta) = (k(\theta), t_1(\theta), \dots, t_n(\theta))$ where $\forall \theta \in \Theta$, $k(\theta) \in K$ and $\sum_i t_i(\theta) \leq 0$.³ For a direct revelation mechanism $\mathcal{D} = ((\Theta_i)_{i \in N}, f(\cdot))$ in this environment, the agent i 's utility function takes the quasi-linear form

$$u_i(x, \theta_i) = u_i(k, t_1, \dots, t_n, \theta_i) = v_i(k, \theta_i) + m_i + t_i,$$

²The reason for the above cyclic dependency is that the preferences do not satisfy the *single peakedness* condition; for more details on the single peakedness condition, the reader may consult Mas-Colell *et al* (1995).

³Note that here we are using symbol k for both as an element of the set K and as a function going from Θ to K . It should be clear from the context as to which of these two we are referring to.

where m_i is agent i 's initial endowment of the money and the function $v_i(\cdot)$ is known as agent i 's valuation function. Recall condition 5 in the definition of mechanism design problem given in § 3. This condition says that the utility function $u_i(\cdot)$ is common knowledge. In the context of a quasi-linear environment, this implies that for any given type θ_i of any agent i , the social planner and every other agent j have a way to figure out the function $v_i(\cdot, \theta_i)$. In many cases, the set Θ_i of the direct revelation mechanism $\mathcal{D} = ((\Theta_i)_{i \in N}, f(\cdot))$ is actually the set of all feasible valuation functions $v_i: K \rightarrow \mathbb{R}$ of agent i . That is, each possible function represents each possible type of agent i . Therefore, in such settings reporting a type is the same as reporting a valuation function.

As far as examples of quasi-linear environment are concerned, all the previously discussed examples, such as fair bilateral trade (Example 3.2), first price auction (Example 3.3), second-price auction (Example 3.4), and generalized Vickrey auction (Example 3.5) are all natural examples of the mechanism in quasi-linear environment.

In the quasi-linear environment, we can define two important properties of a social choice function.

DEFINITION 9.1 (Allocative efficiency (AE)).

We say that SCF $f(\cdot) = (k(\cdot), t_1(\cdot), \dots, t_n(\cdot))$ is allocatively efficient if for each $\theta \in \Theta$, $k(\theta)$ satisfies the following condition⁴

$$k(\theta) \in \arg \max_{k \in K} \sum_{i=1}^n v_i(k, \theta_i). \quad (13)$$

Note that the above definition will make sense only when we ensure that for any given θ , the function $\sum_{i=1}^n v_i(\cdot, \theta_i): K \rightarrow \mathbb{R}$ attains a maximum over the set K . The simplest way to do this is to put a restriction that the function $v_i(\cdot, \theta_i): K \rightarrow \mathbb{R}$ is an upper semi-continuous function for each $\theta_i \in \Theta_i$ and for each $i = 1, \dots, n$.

DEFINITION 9.2 (Budget balance (BB)).

We say that SCF $f(\cdot) = (k(\cdot), t_1(\cdot), \dots, t_n(\cdot))$ is budget balanced if for each $\theta \in \Theta$, $t_1(\theta), \dots, t_n(\theta)$ satisfies the following condition⁵

$$\sum_{i=1}^n t_i(\theta) = 0. \quad (14)$$

The following lemma establishes an important relationship of these two properties of an SCF with the ex-post efficiency of the SCF.

Lemma 9.1. A social choice function $f(\cdot) = (k(\cdot), t_1(\cdot), \dots, t_n(\cdot))$ is ex-post efficient in quasi-linear environment if and only if it is allocatively efficient and budget balanced.

⁴We will keep using the symbol $k^*(\cdot)$ for a function $k(\cdot)$ that satisfies the Equation (13).

⁵Many authors prefer to call this property as strong budget balance and they refer the property of having $\sum_{i=1}^n t_i(\theta) \leq 0$ as weak budget balance. In this thesis, we will use the term budget balance to refer to strong budget balance.

Proof. Let us assume that $f(\cdot) = (k(\cdot), t_1(\cdot), \dots, t_n(\cdot))$ is allocatively efficient and budget balanced. This implies that for any $\theta \in \Theta$, we have

$$\begin{aligned} \sum_{i=1}^n u_i(f(\theta), \theta_i) &= \sum_{i=1}^n v_i(k(\theta), \theta_i) + \sum_{i=1}^n t_i(\theta) \\ &= \sum_{i=1}^n v_i(k(\theta), \theta_i) + 0 \\ &\geq \sum_{i=1}^n v_i(k, \theta_i) + \sum_{i=1}^n t_i; \quad \forall x = (k, t_1, \dots, t_n) \\ &= \sum_{i=1}^n u_i(x, \theta_i); \quad \forall (k, t_1, \dots, t_n) \in X. \end{aligned}$$

That is, if the SCF is allocatively efficient and budget balanced then for any type profile θ of the agent the outcome chosen by the social choice function will be such that it maximizes the total utility derived by all the agents. This will automatically imply that the SCF is ex-post efficient.

To prove the other part, we will first show that if $f(\cdot)$ is not allocatively efficient then it cannot be ex-post efficient and next we will show that if $f(\cdot)$ is not budget balanced then it cannot be ex-post efficient. These two facts together will imply that if $f(\cdot)$ is ex-post efficient then it will have to be allocatively efficient and budget balanced, thus completing the proof of the lemma.

To start with, let us assume that $f(\cdot)$ is not allocatively efficient. This means that $\exists \theta \in \Theta$, and $k \in K$ such that

$$\sum_{i=1}^n v_i(k, \theta_i) > \sum_{i=1}^n v_i(k(\theta), \theta_i).$$

This implies that there exists at least one agent j for whom $v_j(k, \theta_j) > v_j(k(\theta), \theta_j)$. Now consider the following alternative x

$$x = (k, (t_i = t_i(\theta) + v_i(k(\theta), \theta_i) - v_i(k, \theta_i))_{i \neq j}, t_j = t_j(\theta)).$$

It is easy to verify that $u_i(x, \theta_i) = u_i(f(\theta), \theta_i) \quad \forall i \neq j$ and $u_j(x, \theta_j) > u_j(f(\theta), \theta_j)$ implying that $f(\cdot)$ is not ex-post efficient.

Next, we assume that $f(\cdot)$ is not budget balanced. This means that there exists at least one agent j for whom $t_j(\theta) < 0$. Let us consider the following alternative x

$$x = (k, (t_i = t_i(\theta))_{i \neq j}, t_j = 0).$$

It is easy to verify that for the above alternative x , we have $u_i(x, \theta_i) = u_i(f(\theta), \theta_i) \quad \forall i \neq j$ and $u_j(x, \theta_j) > u_j(f(\theta), \theta_j)$ implying that $f(\cdot)$ is not ex-post efficient. Q.E.D.

The next lemma summarizes another fact about the social choice function in quasi-linear environment.

Lemma 9.2. All social choice functions in quasi-linear environment are non-dictatorial.

Proof. If possible assume that SCF $f(\cdot)$ is dictatorial. This means that for each $\theta \in \Theta$, we have

$$u_d(f(\theta), \theta_d) \geq u_d(x, \theta_d) \quad \forall x \in X,$$

where d is a dictatorial agent. However, because of quasi-linear environment, we have $u_d(f(\theta), \theta_d) = v_d(k(\theta), \theta_d) + t_d(\theta)$. Now consider the following alternative $x \in X$:

$$x = \begin{cases} (k(\theta), (t_i = t_i(\theta))_{i \neq d}, t_d = t_d(\theta) - \sum_{i=1}^n t_i(\theta)) : \sum_{i=1}^n t_i(\theta) < 0 \\ (k(\theta), (t_i = t_i(\theta))_{i \neq d, j}, t_d = t_d(\theta) + \epsilon, t_j = t_j(\theta) - \epsilon) : \sum_{i=1}^n t_i(\theta) = 0, \end{cases}$$

where $\epsilon > 0$ is any arbitrary number and j is any agent other than d . It is easy to verify that for the above constructed x , we have $u_d(x, \theta_d) > u_d(f(\theta), \theta_d)$ contradicting the assumption that $f(\cdot)$ is dictatorial. Q.E.D.

In view of Lemma 9.2, the social planner need not have to worry about the non-dictatorial property of the social choice function in quasi-linear environments and he can simply look for whether there exists any SCF which is both ex-post efficient and dominant strategy incentive compatible. Further in the light of Lemma 9.1, we can say that the social planner can look for an SCF which is allocatively efficient, budget balanced, and dominant strategy incentive compatible. Once again, the question arises whether there exists any such social choice function which satisfies all these three properties — AE, BB, and DSIC. We explore the answer to this question in what follows.

10. Groves mechanisms

The following theorem, due to Groves (1973) confirms that in quasi-linear environment, there exist social choice functions which are both allocatively efficient and truthfully implementable in dominant strategies (dominant strategy incentive compatible).

Theorem 10.1 (Groves' theorem). *Let the social choice function $f(\cdot) = (k^*(\cdot), t_1(\cdot), \dots, t_n(\cdot))$ be allocatively efficient. This function can be truthfully implemented in dominant strategies if it satisfies the following payment structure (popularly known as Groves payment (incentive) scheme):*

$$t_i(\theta) = \left[\sum_{j \neq i} v_j(k^*(\theta), \theta_j) \right] + h_i(\theta_{-i}) \quad \forall i = 1, \dots, n \quad (15)$$

where $h_i(\cdot)$ is any arbitrary function of θ_{-i} up to satisfying the feasibility condition $\sum_i t_i(\theta) \leq 0 \quad \forall \theta \in \Theta$.

For proof of the Groves theorem, refer to Proposition 23.C.4 of Mas-Colell *et al* (1995). Following are a few interesting implications of the above theorem.

- (i) Given the announcements θ_{-i} of agents $j \neq i$, agent i 's transfer depends on his announced type only through his announcement's effect on the project choice $k^*(\theta)$.

- (ii) The change in the monetary transfer of agent i when his type changes from θ_i to $\hat{\theta}_i$ is equal to the effect that the corresponding change in project choice has on total value of the rest of the agents. That is,

$$t_i(\theta_i, \theta_{-i}) - t_i(\hat{\theta}_i, \theta_{-i}) = \sum_{j \neq i} \left[v_j(k^*(\theta_i, \theta_{-i}), \theta_j) - v_j(k^*(\hat{\theta}_i, \theta_{-i}), \theta_j) \right].$$

Another way of describing this is to say that the change in monetary transfer to agent i reflects exactly the externality he is imposing on the other agents.

After the famous result of Groves, a direct revelation mechanism in which the implemented SCF is allocatively efficient and satisfies the Groves payment scheme is called as *Groves Mechanism*.

DEFINITION 10.1 (Groves mechanisms).

A direct revelation mechanism, $\mathcal{D} = ((\Theta_i)_{i \in N}, f(\cdot))$ in which $f(\cdot) = (k(\cdot), t_1(\cdot), \dots, t_n(\cdot))$ satisfies (13) and (15) is known as Groves mechanism.⁶

In practice, Groves mechanisms are popularly known as Vickrey–Clarke–Groves (VCG) mechanisms because Clarke mechanism is a special case of Groves mechanism and Vickrey mechanism is a special case of Clarke mechanism. We will discuss this relationship later in this paper.

Groves theorem provides a sufficiency condition under which an allocatively efficient (AE) SCF will be DSIC. The following theorem due to Green and Laffont (1979) provides a set of conditions under which the condition of Groves theorem also becomes a necessary condition for an AE SCF to be DSIC. In this theorem, we let \mathcal{F} denote the set of all possible functions $f: K \rightarrow \mathbb{R}$.

Theorem 10.2 (First characterization theorem of Green–Laffont). *Suppose that for each agent $i = 1, \dots, n$ $\{v_i(\cdot, \theta_i): \theta_i \in \Theta_i\} = \mathcal{F}$; that is, every possible valuation function from K to \mathbb{R} arises for some $\theta_i \in \Theta_i$. Then any allocatively efficient (AE) social choice function $f(\cdot)$ will be dominant strategy incentive compatible (DSIC) if and only if it satisfies the Groves payment scheme given by (15).*

Note that in the above theorem, every possible valuation function from K to \mathbb{R} arises for some $\theta_i \in \Theta_i$. Therefore, in many cases, depending upon the structure of the compact set \mathcal{K} , it is quite possible that for some type profile $\theta = (\theta_1, \dots, \theta_n)$, the maximum of the function $\sum_{i=1}^n v_i(\cdot, \theta_i)$ over the set \mathcal{K} may not exist. In such cases, the set of AE social choice functions would be empty and the theorem will make no sense. One possible way to get away with this difficulty is to assume that the set \mathcal{K} is a finite set. Another solution is to restrict the allowable valuation functions to the class of continuous functions. The following characterization theorem of Green and Laffont (1979) fixes this problem by replacing the \mathcal{F} with \mathcal{F}_c where \mathcal{F}_c denotes the set of all possible continuous functions $f: K \rightarrow \mathbb{R}$.

Theorem 10.3 (Second characterization theorem of Green–Laffont). *Suppose that for each agent $i = 1, \dots, n$ $\{v_i(\cdot, \theta_i): \theta_i \in \Theta_i\} = \mathcal{F}_c$; that is, every possible continuous valuation function from K to \mathbb{R} arises for some $\theta_i \in \Theta_i$. Then any allocatively efficient (AE) social*

⁶We will sometimes abuse the terminology and simply refer to a SCF $f(\cdot)$ satisfying (13) and (15) as Groves mechanism.

choice function $f(\cdot)$ will be dominant strategy incentive compatible (DSIC) if and only if it satisfies the Groves payment scheme given by (15).

10.1 Groves mechanisms and budget balance

Note that a Groves mechanism always satisfies the properties of AE and DSIC. Therefore, if a Groves mechanism is budget balanced then it will solve the problem of the social planner because it will then be ex-post efficient and dominant strategy incentive compatible. By looking at the definition of the Groves mechanism, one can conclude that it is the functions $h_i(\cdot)$ that decide whether or not the Groves mechanism is budget balanced. The natural question that arises now is whether there exists a way of defining functions $h_i(\cdot)$ such that the Groves mechanism is budget balanced. In what follows, we present one possibility and one impossibility result in this regard.

10.2 Possibility and impossibility results for quasi-linear environments

Green and Laffont (1979) showed that in quasi-linear environment, if the set of possible types for each agent is sufficiently rich then ex-post efficiency and DSIC cannot be achieved together. The precise statement is given in the form of following theorem.

Theorem 10.4 (Green–Laffont impossibility theorem). *Suppose that for each agent $i = 1, \dots, n$, $\{v_i(\cdot, \theta_i): \theta_i \in \Theta_i\} = \mathcal{F}$; that is, every possible valuation function from K to \mathbb{R} arises for some $\theta_i \in \Theta_i$. Then there is no social choice function which is ex-post efficient and DSIC.*

Thus, the above theorem says that if the set of possible types for each agent is sufficiently rich then there is no hope of finding a way to define the functions $h_i(\cdot)$ in Groves payment scheme so that we have $\sum_{i=1}^n t_i(\theta) = 0$. However, one special case in which a more positive result does obtain is summarized in the form of following possibility result.

Theorem 10.5 (Possibility result for budget balance of Groves mechanisms). *If there is at least one agent whose preferences are known (i.e. his type set is a singleton set) then it is possible to choose the functions $h_i(\cdot)$ so that $\sum_{i=1}^n t_i(\theta) = 0$.*

Proof. Let agent i be such that his preferences are known, that is $\Theta_i = \{\theta_i\}$. In view of this condition, it is easy to see that for an allocatively efficient social choice function $f(\cdot) = (k^*(\cdot), t_1(\cdot), \dots, t_n(\cdot))$, the allocation $k^*(\cdot)$ depends only on the types of the agents other than i . That is, the allocation $k^*(\cdot)$ is a mapping from Θ_{-i} to K . Let us define the functions $h_j(\cdot)$ in the following manner.

$$h_j(\theta_{-j}) = \begin{cases} h_j(\theta_{-j}) : j \neq i \\ -\sum_{r \neq i} h_r(\theta_{-r}) - (n-1) \sum_{r=1}^n v_r(k^*(\theta), \theta_r) : j = i. \end{cases}$$

It is easy to see that under the above definition of the functions $h_i(\cdot)$, we will have $t_i(\theta) = -\sum_{j \neq i} t_j(\theta)$.

10.3 Clarke (pivotal) mechanisms: special case of Groves mechanisms

A special case of Groves mechanism was discovered independently by Clarke (1971) and is known as *Clarke*, or *pivotal* mechanism. It is a special case of Groves mechanism in the sense

of using a particular form for the function $h_i(\cdot)$. In the Clarke mechanism, the function $h_i(\cdot)$ is given by the following relation:

$$h_i(\theta_{-i}) = - \sum_{j \neq i} v_j(k_{-i}^*(\theta_{-i}), \theta_j) \quad \forall \theta_{-i} \in \Theta_{-i}, \forall i = 1, \dots, n, \quad (16)$$

where $k_{-i}^*(\theta_{-i}) \in K_{-i}$ is the choice of a project which is allocatively efficient if there were only the $n - 1$ agents $j \neq i$. Formally, $k_{-i}^*(\theta_{-i})$ must satisfy the following condition.

$$\sum_{j \neq i} v_j(k_{-i}^*(\theta_{-i}), \theta_j) \geq \sum_{j \neq i} v_j(k, \theta_j) \quad \forall k \in K_{-i}, \quad (17)$$

where the set K_{-i} is the set of project choices available when agent i is absent. Substituting the value of $h_i(\cdot)$ from Equation (16) in Equation (15), we get the following expression for agent i 's transfer in the Clarke mechanism

$$t_i(\theta) = \left[\sum_{j \neq i} v_j(k^*(\theta), \theta_j) \right] - \left[\sum_{j \neq i} v_j(k_{-i}^*(\theta_{-i}), \theta_j) \right]. \quad (18)$$

10.4 Clarke mechanisms and weak budget balance

Recall from the definition of Groves mechanisms that, for weak budget balance, we should choose the functions $h_i(\theta_{-i})$ in such a way that the weak budget balance condition $\sum_{i=1}^n t_i(\theta) \leq 0$ is satisfied. In this sense, the Clarke mechanism is a useful special-case because it achieves weak budget balance under fairly general settings. In order to understand these general sufficiency conditions, we define following quantities.

$$B^*(\theta) = \left\{ k \in K \mid k \in \arg \max_{k \in K} \sum_{j=1}^n v_j(k, \theta_j) \right\}$$

$$B^*(\theta_{-i}) = \left\{ k \in K_{-i} \mid k \in \arg \max_{k \in K_{-i}} \sum_{j \neq i} v_j(k, \theta_j) \right\},$$

where $B^*(\theta)$ is the set of project choices that are allocatively efficient when all the agents are present there in the system. Similarly, $B^*(\theta_{-i})$ is the set of project choices that are allocatively efficient if there were $n - 1$ agents $j \neq i$. It is obvious that $k^*(\theta) \in B^*(\theta)$ and $k_{-i}^*(\theta_{-i}) \in B^*(\theta_{-i})$.

Using the above quantities, we define the following properties of a direct revelation mechanism in quasi-linear environment.

DEFINITION 10.2 (No single agent effect).

We say that mechanism \mathcal{M} has no single agent effect if for each agent i , each $\theta \in \Theta$, and each $k^*(\theta) \in B^*(\theta)$, we have a $k \in K_{-i}$ such that

$$\sum_{j \neq i} v_j(k, \theta_j) \geq \sum_{j \neq i} v_j(k^*(\theta), \theta_j).$$

In view of the above properties, we have the following proposition that gives sufficiency condition for Clarke mechanism to be weak budget balanced.

PROPOSITION 10.1.

If the Clarke mechanism has no single agent effect, then transfer of each agent would be non-positive, that is, $t_i(\theta_i) \leq 0 \forall \theta \in \Theta; \forall i = 1, \dots, n$. In such a situation, the Clarke mechanism would satisfy the weak budget balance property.

Proof. Note that by virtue of no single agent effect, for each agent i , each $\theta \in \Theta$, and each $k^*(\theta) \in B^*(\theta)$, $\exists k \in K_{-i}$ such that

$$\sum_{j \neq i} v_j(k, \theta_j) \geq \sum_{j \neq i} v_j(k^*(\theta), \theta_j).$$

However, by definition of $k_{-i}^*(\theta_{-i})$, given by Equation (17), we have

$$\sum_{j \neq i} v_j(k_{-i}^*(\theta_{-i}), \theta_j) \geq \sum_{j \neq i} v_j(k, \theta_j) \forall k \in K_{-i}.$$

Combining the above two facts, we get

$$\begin{aligned} \sum_{j \neq i} v_j(k_{-i}^*(\theta_{-i}), \theta_j) &\geq \sum_{j \neq i} v_j(k^*(\theta), \theta_j) \\ &\Rightarrow 0 \geq t_i(\theta) \\ &\Rightarrow 0 \geq \sum_{i=1}^n t_i(\theta). \end{aligned}$$

Q.E.D.

Now we can make following assertions with respect to these definitions. In what follows is an interesting corollary of the above proposition.

COROLLARY 10.1.

- (i) $t_i(\theta) = 0$ if $k^*(\theta) \in B^*(\theta_{-i})$. That is, agent i 's monetary transfer is zero if his announcement does not change the project decision relative to what would be allocatively efficient for agents $j \neq i$ in isolation.
- (ii) $t_i(\theta) < 0$ if $k^*(\theta) \notin B^*(\theta_{-i})$. That is, agent i 's monetary transfer is negative if his announcement changes the project decision relative to what would be allocatively efficient for agents $j \neq i$ in isolation. In such situation, the agent i is known to be 'pivotal' to the efficient project choice and he pays a tax equal to his effect on the other agents.

The following picture summarizes the conclusions of this section by showing how the space of social choice functions looks like in quasi-linear environment. In what follows we discuss a few examples of direct revelation mechanisms in quasi-linear environment and study the various properties of the underlying SCF.

10.5 Fair bilateral trade is $BB + AE$ but not DSIC

Consider the fair bilateral trade Example 3.2 in which the good is always allocated to the agent who values it the most. Therefore, the SCF is allocatively efficient. The SCF here is

10.7 Second-price sealed bid (Vickrey) auction is AE + DSIC but not BB

Consider Example 3.4 of second-price (Vickrey) auction. Once again we assume that $S_i = \Theta_i = [\underline{\theta}_i, \bar{\theta}_i] \quad \forall i \in N$. In such a case, the second-price auction becomes a direct revelation mechanism $\mathcal{D} = ((\Theta_i)_{i \in N}, f(\cdot))$, where $f(\cdot)$ is an SCF which is same as outcome rule of the second-price auction. We have already shown that Vickrey auction is a special case of GVA in which the auctioneer is selling just single unit of a single item. Given that the Vickrey auction a special case of the GVA, we can assert that the SCF $f(\cdot)$ used by the auctioneer in Vickrey auction is AE + DSIC but not BB because the SCF used in GVA is also AE + DSIC and not BB, which we are going to show next.

10.8 Generalized Vickrey auction is AE + DSIC but not BB

Consider Example 3.5 of generalized Vickrey auction. Let us assume that $S_i = \Theta_i = [\underline{V}_i, \bar{V}_i]^{(2^m - 1)} \quad \forall i \in N$. In such a case, the GVA auction becomes a direct revelation mechanism $\mathcal{D} = ((\Theta_i)_{i \in N}, f(\cdot))$, where $f(\cdot)$ is an SCF which is same as outcome rule of the GVA auction. It is easy to see that the under this SCF $f(\cdot)$, the bundles are allocated in such a way that total value of all the agents gets maximized. This implies that the SCF in this example is allocatively efficient. Also, note that the payment structure is the same as the Clarke mechanism. Thus, the direct revelation mechanism used by the auctioneer in GVA is a Clarke mechanism. Therefore, the SCF is automatically DSIC. However, note that SCF certainly not budget balanced because we are not including the auctioneer into the system. If auctioneer is also part of the system then by virtue of theorem 10.5, we can claim that the SCF would be BB also.

Table 1 summarizes the properties of the SCFs discussed in above examples. Figure 9 summarizes the relationship among various SCFs discussed in above examples.

11. Bayesian implementation

Recall that we mentioned two possible routes to get around the Gibbard–Satterthwaite impossibility theorem. The first was to focus on restricted environments like quasi-linear environment, and the second one was to weaken the implementation concept and look for an SCF which is ex-post efficient, non-dictatorial, and Bayesian incentive compatible. In this section, our objective is to explore the second route.

Throughout this section, we will once again be working within the quasi-linear environment. As we saw earlier, the quasi-linear environments have a nice property that every social choice function in these environments is non-dictatorial. Therefore, while working within quasi-linear environment, we do not have to worry about non-dictatorial part of the social choice function. We can just investigate whether there exists any SCF in quasi-linear environment, which is both *ex-post efficient* and *BIC*, or equivalently which has three

Table 1. Properties of social choice functions in quasi-linear environment.

SCF	AE	BB	DSIC
Fair bilateral trade	✓	✓	×
First-price auction	✓	×	×
Vickrey auction	✓	×	✓
GVA	✓	×	✓

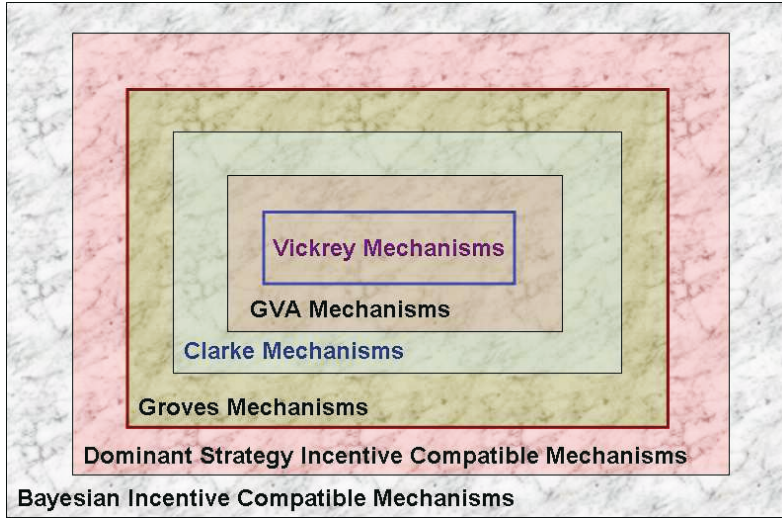


Figure 9. Space of BIC and DSIC social choice functions in quasi-linear environment.

properties — *AE*, *BB*, and *BIC*. Recall that in the previous section, we have already addressed the question whether there exists any SCF in quasi-linear environment which is *AE*, *BB*, and *DSIC* and we found that hardly any function satisfies all these three properties. On the contrary, in this section, we will show that a wide range of SCF in quasi-linear environment satisfy three properties — *AE*, *BB*, and *BIC*.

11.1 (*AE + BB + BIC*) is possible: expected externality mechanisms

The following theorem, due to d'Aspremont and Gérard-Varet (1979) and Arrow (1979) confirms that in quasi-linear environment, there exist social choice functions which are both ex-post efficient (allocatively efficient + budget balance) and truthfully implementable in Bayesian–Nash equilibrium (Bayesian incentive compatible). We refer this theorem by *dAGVA* theorem.

Theorem 11.1 (The dAGVA theorem). *Let the social choice function $f(\cdot) = (k^*(\cdot), t_1(\cdot), \dots, t_n(\cdot))$ be allocatively efficient and the agents' types be statistically independent of each other (i.e. the density $\phi(\cdot)$ has the form $\phi_1(\cdot) \times \dots \times \phi_n(\cdot)$). This function can be truthfully implemented in Bayesian–Nash equilibrium if it satisfies the following payment structure, popularly known as dAGVA payment (incentive) scheme.*

$$t_i(\theta) = E_{\tilde{\theta}_{-i}} \left[\sum_{j \neq i} v_j(k^*(\theta_i, \tilde{\theta}_{-i}), \tilde{\theta}_j) \right] + h_i(\theta_{-i}) \quad \forall i = 1, \dots, n; \quad \forall \theta \in \Theta, \quad (19)$$

where $h_i(\cdot)$ is any arbitrary function of θ_{-i} . Moreover, it is always possible to choose the functions $h_i(\cdot)$ such that $\sum_{i=1}^n t_i(\theta) = 0$.

Proof. Let the social choice function $f(\cdot) = (k^*(\cdot), t_1(\cdot), \dots, t_n(\cdot))$ be allocatively efficient, i.e. it satisfies the condition (13), and also satisfies the dAGVA payment scheme (19). Consider

$$E_{\theta_{-i}} [u_i(f(\theta_i, \theta_{-i}), \theta_i) | \theta_i] = E_{\theta_{-i}} [v_i(k^*(\theta_i, \theta_{-i}), \theta_i) + t_i(\theta_i, \theta_{-i}) | \theta_i].$$

Since θ_i and θ_{-i} are statistically independent, the expectation can be taken without conditioning on θ_i . This will give us

$$\begin{aligned} & E_{\theta_{-i}} [u_i(f(\theta_i, \theta_{-i}), \theta_i) | \theta_i] \\ &= E_{\theta_{-i}} \left[v_i(k^*(\theta_i, \theta_{-i}), \theta_i) + h_i(\theta_{-i}) + E_{\tilde{\theta}_{-i}} \left[\sum_{j \neq i} v_j(k^*(\theta_i, \tilde{\theta}_{-i}), \tilde{\theta}_j) \right] \right] \\ &= E_{\theta_{-i}} \left[\sum_{j=1}^n v_j(k^*(\theta_i, \theta_{-i}), \theta_j) \right] + E_{\theta_{-i}} [h_i(\theta_{-i})]. \end{aligned}$$

Since $k^*(\cdot)$ satisfies the condition (13),

$$\sum_{j=1}^n v_j(k^*(\theta_i, \theta_{-i}), \theta_j) \geq \sum_{j=1}^n v_j(k^*(\hat{\theta}_i, \theta_{-i}), \theta_j) \quad \forall \hat{\theta}_i \in \Theta_i.$$

Thus, we get

$$\begin{aligned} & E_{\theta_{-i}} \left[\sum_{j=1}^n v_j(k^*(\theta_i, \theta_{-i}), \theta_j) \right] + E_{\theta_{-i}} [h_i(\theta_{-i})] \\ & \geq E_{\theta_{-i}} \left[\sum_{j=1}^n v_j(k^*(\hat{\theta}_i, \theta_{-i}), \theta_j) \right] + E_{\theta_{-i}} [h_i(\theta_{-i})] \quad \forall \hat{\theta}_i \in \Theta_i. \end{aligned}$$

Again by making use of statistical independence we can rewrite the above inequality in the following form

$$E_{\theta_{-i}} [u_i(f(\theta_i, \theta_{-i}), \theta_i) | \theta_i] \geq E_{\theta_{-i}} [u_i(f(\hat{\theta}_i, \theta_{-i}), \theta_i) | \theta_i] \quad \forall \hat{\theta}_i \in \Theta_i.$$

This shows that when agents $j \neq i$ announce their types truthfully, agent i finds the truth telling is his optimal strategy, thus proving that the SCF is BIC. We now show that the functions $h_i(\cdot)$ can be chosen to guarantee $\sum_{i=1}^n t_i(\theta) = 0$. Let us define,

$$\begin{aligned} \xi_i(\theta_i) &= E_{\tilde{\theta}_{-i}} \left[\sum_{j \neq i} v_j(k^*(\theta_i, \tilde{\theta}_{-i}), \tilde{\theta}_j) \right] \quad \forall i = 1, \dots, n \\ h_i(\theta_{-i}) &= - \left(\frac{1}{n-1} \right) \sum_{j \neq i} \xi_j(\theta_j) \quad \forall i = 1, \dots, n. \end{aligned}$$

In view of the above definitions, we can say that

$$\begin{aligned} t_i(\theta) &= \xi_i(\theta_i) - \left(\frac{1}{n-1} \right) \sum_{j \neq i} \xi_j(\theta_j) \\ \Rightarrow \sum_{i=1}^n t_i(\theta) &= \sum_{i=1}^n \xi_i(\theta_i) - \left(\frac{1}{n-1} \right) \sum_{i=1}^n \sum_{j \neq i} \xi_j(\theta_j) \end{aligned}$$

$$\begin{aligned} \Rightarrow \sum_{i=1}^n t_i(\theta) &= \sum_{i=1}^n \xi_i(\theta_i) - \left(\frac{1}{n-1}\right) \sum_{i=1}^n (n-1)\xi_j(\theta_j) \\ \Rightarrow \sum_{i=1}^n t_i(\theta) &= 0. \end{aligned}$$

Q.E.D.

The budget balanced payment structure of the agents in the above mechanism can be given a nice graph theoretic interpretation. Imagine a directed graph $G = (V, A)$ where V is the set of $n + 1$ vertices, numbered $0, 1, \dots, n$, and A is the set of $[n + n(n - 1)]$ directed arcs. The vertices starting from 1 through n correspond to the n agents involved into the system and the vertex number 0 corresponds to the social planner. The set A consists of two types of the directed arcs:

- (i) Arcs $0 \rightarrow i \quad \forall i = 1, \dots, n$
- (ii) Arcs $i \rightarrow j \quad \forall i, j \in \{1, 2, \dots, n\}; i \neq j$.

Each of the arcs $0 \rightarrow i$ carries a flow of $t_i(\theta)$ and each of the arcs $i \rightarrow j$ carries a flow of $\frac{\xi_i(\theta_i)}{n-1}$. Thus the total outflow from a node $i \in \{1, 2, \dots, n\}$ is $\xi_i(\theta_i)$ and total inflow to the node i from nodes $j \in \{1, 2, \dots, n\}$ is $-h_i(\theta_{-i}) = \left(\frac{1}{n-1}\right) \sum_{j \neq i} \xi_j(\theta_j)$. Thus for any node i , $t_i(\theta) + h_i(\theta_{-i})$ is the net outflow which it is receiving from node 0 in order to respect the flow conservation constraint. Thus, if $t_i(\cdot)$ is positive then the agent i receives the money from the social planner and if it is negative, then the agent pays the money to the social planner. However, by looking at flow conservation equation for node 0, we can say that total payment received by the planner from the agents and total payment made by the planner to the agents will add up to zero. In graph theoretic interpretation, the flow from node i to node j can be justified as follows. Each agent i first evaluates the expected total valuation that would be generated together by all his rival agents in his absence, which turns out to be $\xi_i(\theta_i)$. Now, agent i divides it equally among the rival agents and pays to every rival agent an amount equivalent to this. The idea can be better understood with the help of figure 10 which depicts the three agents case.

After the results of d'Aspremont and Gérard-Varet (1979) and Arrow (1979), a direct revelation mechanism in which SCF is allocatively efficient and satisfies the dAGVA payment scheme is called as *dAGVA mechanism/expected externality mechanism/expected Groves mechanism*.

DEFINITION 11.1 (dAGVA/expected externality/expected Groves Mechanisms).

A direct revelation mechanism, $\mathcal{D} = ((\Theta_i)_{i \in N}, f(\cdot))$ in which $f(\cdot) = (k(\cdot), t_1(\cdot), \dots, t_n(\cdot))$ satisfies (13) and (19) is known as dAGVA/expected externality/expected Groves Mechanism.⁷

In view of the definition of dAGVA mechanisms, the figure 8 can be enriched by including the space of dAGVA mechanisms. This is shown in figure 11.

⁷We will sometimes abuse the terminology and simply refer to a SCF $f(\cdot)$ satisfying (13) and (19) as dAGVA/expected externality/expected Groves Mechanisms.

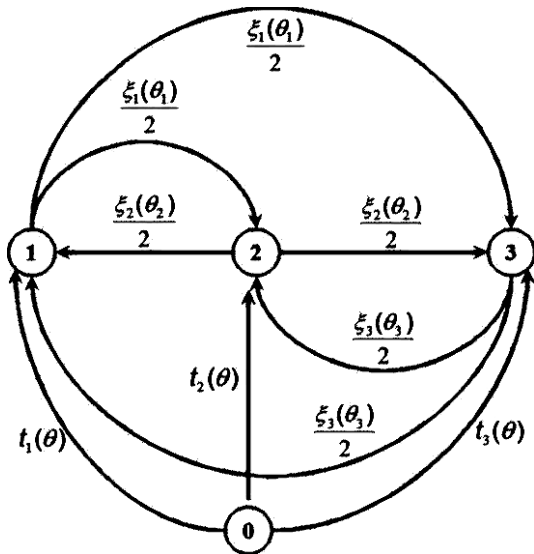


Figure 10. Payment structure in budget balance expected externality mechanism.

11.2 BIC in linear environment

The linear environment is a special, but often-studied, subclass of quasi-linear environment. This environment is a restricted version of the quasi-linear environment in following sense.

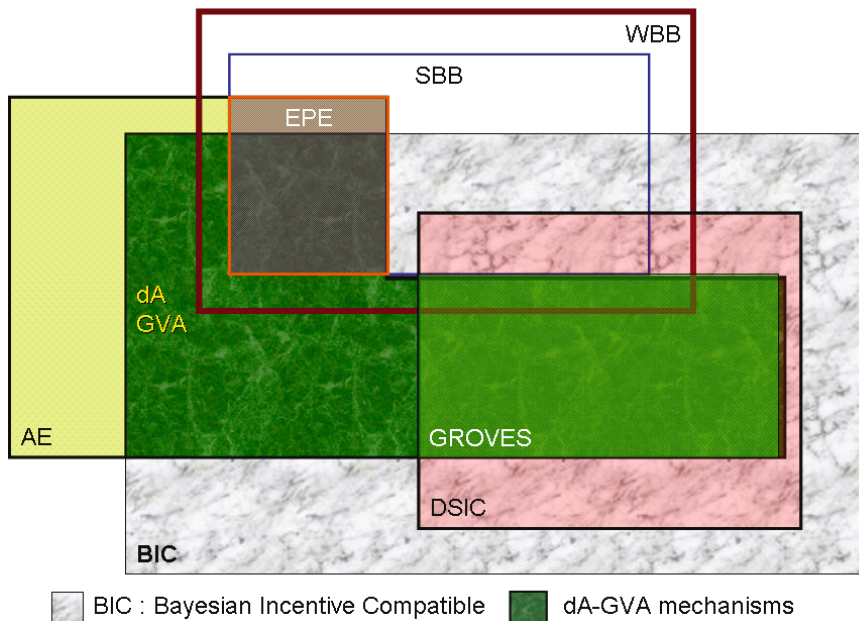


Figure 11. Space of social choice functions in quasi-linear environment.

- (i) Each agent i 's type lies in an interval $\Theta_i = [\underline{\theta}_i, \bar{\theta}_i] \subset \mathbb{R}$ with $\underline{\theta}_i < \bar{\theta}_i$
- (ii) Agents' types are statistically independent, that is, the density $\phi(\cdot)$ has the form $\phi_1(\cdot) \times \cdots \times \phi_n(\cdot)$
- (iii) $\phi_i(\theta_i) > 0 \forall \theta_i \in [\underline{\theta}_i, \bar{\theta}_i] \forall i = 1, \dots, n$
- (iv) Each agent i 's utility function takes the following form

$$u_i(x, \theta_i) = \theta_i v_i(k) + m_i + t_i.$$

The linear environment has very interesting properties in terms of characterizing the BIC social choice functions. Before we present Myerson's characterization theorem for BIC social choice functions in linear environment, we would like to define following quantities with regard to any social choice function $f(\cdot) = (k(\cdot), t_1(\cdot), \dots, t_n(\cdot))$ in this environment.

- Let $\bar{t}_i(\hat{\theta}_i) = E_{\theta_{-i}}[t_i(\hat{\theta}_i, \theta_{-i})]$ be agent i 's expected transfer given that he announces his type to be $\hat{\theta}_i$ and that all agents $j \neq i$ truthfully reveal their types.
- Let $\bar{v}_i(\hat{\theta}_i) = E_{\theta_{-i}}[v_i(\hat{\theta}_i, \theta_{-i})]$ be agent i 's expected 'benefits' given that he announces his type to be $\hat{\theta}_i$ and that all agents $j \neq i$ truthfully reveal their types.
- Let $U_i(\hat{\theta}_i|\theta_i) = E_{\theta_{-i}}[u_i(f(\hat{\theta}_i, \theta_{-i}), \theta_i)|\theta_i]$ be agent i 's expected utility when his type is θ_i , he announces his type to be $\hat{\theta}_i$, and that all agents $j \neq i$ truthfully reveal their types. It is easy to verify from previous two definitions that

$$U_i(\hat{\theta}_i|\theta_i) = \theta_i \bar{v}_i(\hat{\theta}_i) + \bar{t}_i(\hat{\theta}_i).$$

- Let $U_i(\theta_i) = U_i(\theta_i|\theta_i)$ be the agent i 's expected utility conditional on his type being θ_i when he and all other agents report their true types. It is easy to verify that

$$U_i(\theta_i) = \theta_i \bar{v}_i(\theta_i) + \bar{t}_i(\theta_i).$$

In view of the paraphernalia developed, we now present Myerson's (1981) theorem for characterizing the BIC social choice functions in this environment.

Theorem 11.2 (Myerson's characterization theorem). *In linear environment, a social choice function $f(\cdot) = (k(\cdot), t_1(\cdot), \dots, t_n(\cdot))$ is BIC if and only if, for all $i = 1, \dots, n$,*

- (i) $\bar{v}_i(\cdot)$ is non-decreasing
- (ii) $U_i(\theta_i) = U_i(\underline{\theta}_i) + \int_{\underline{\theta}_i}^{\theta_i} \bar{v}_i(s) ds \quad \forall \theta_i.$

For proof of the above theorem, refer to Proposition 23-D-2 of Mas-Colell *et al* (1995). The above theorem shows that to identify all BIC social choice functions in linear environment, we can proceed as follows: First, identify which functions $k(\cdot)$ lead every agent i 's expected benefit function $\bar{v}_i(\cdot)$ to be non-decreasing. Then, for each such function identify transfer functions $\bar{t}_1(\cdot), \dots, \bar{t}_n(\cdot)$ that satisfy the second condition of the above proposition. Substituting for $U_i(\cdot)$ in the second condition above, we get that expected transfer functions are precisely those which satisfy, for $i = 1, \dots, n$,

$$\bar{t}_i(\theta_i) = \bar{t}_i(\underline{\theta}_i) + \underline{\theta}_i \bar{v}_i(\underline{\theta}_i) - \theta_i \bar{v}_i(\theta_i) + \int_{\underline{\theta}_i}^{\theta_i} \bar{v}_i(s) ds,$$

for some constant $\bar{t}_i(\theta_i)$. Finally, choose any set of transfer functions $t_1(\cdot), \dots, t_n(\cdot)$ such that $E_{\theta_{-i}}[t_i(\theta_i, \theta_{-i})] = \bar{t}_i(\theta_i)$ for all θ_i . In general, there are many such functions, $t_i(\cdot, \cdot)$; one, for example, is simply $t_i(\theta_i, \theta_{-i}) = \bar{t}_i(\theta_i)$.

In what follows we discuss a few examples where the environment is linear and analyse the BIC property of the social choice function by means of Myerson's characterization theorem.

11.3 Fair bilateral trade in linear environment

Once again consider the Example 3.2 of fair bilateral trade. Recall that each agent i 's type lies in an interval $\Theta_i = [\underline{\theta}_i, \bar{\theta}_i]$. Let us impose the additional conditions on the environment to make it linear. We assume that

- (i) Agents' types are statistically independent, that is, the density $\phi(\cdot)$ has the form $\phi_1(\cdot) \times \phi_2(\cdot)$
- (ii) Let each agent i draw his type from the set $[\underline{\theta}_i, \bar{\theta}_i]$ by means of a uniform distribution, that is $\phi_i(\theta_i) = 1/(\bar{\theta}_i - \underline{\theta}_i) \forall \theta_i \in [\underline{\theta}_i, \bar{\theta}_i] \forall i = 1, 2$.

Note that the utility function of the agents in this example are given by

$$u_i(f(\theta), \theta_i) = \theta_i y_i(\theta) + t_i(\theta) \forall i = 1, 2.$$

Thus, viewing $y_i(\theta) = v_i(k(\theta))$ will confirm that these utility functions also satisfy the fourth condition required for linear environment. Now we can apply Myerson's characterization theorem to test the Bayesian incentive compatibility of the SCF involved here. It is easy to see that $\bar{v}_1(\theta_1) = \bar{y}_1(\theta_1) = 1 - \Phi_2(\theta_1)$ is not a non-decreasing function. Therefore, we can claim that fair bilateral trade is not BIC.

11.4 First-price sealed bid auction in linear environment

Once again, consider the Example 3.3 of first-price sealed bid auction. Let us assume that $S_i = \Theta_i = [\underline{\theta}_i, \bar{\theta}_i] \forall i \in N$. In such a case, the first-price auction becomes a direct revelation mechanism $\mathcal{D} = ((\Theta_i)_{i \in N}, f(\cdot))$, where $f(\cdot)$ is an SCF which is same as outcome rule of the first-price auction. Let us impose the additional conditions on the environment to make it linear. We assume that

- (i) Bidders' types are statistically independent, that is, the density $\phi(\cdot)$ has the form $\phi_1(\cdot) \times \dots \times \phi_n(\cdot)$
- (ii) Let each bidder draw his type from the set $[\underline{\theta}_i, \bar{\theta}_i]$ by means of a uniform distribution, that is $\phi_i(\theta_i) = 1/(\bar{\theta}_i - \underline{\theta}_i) \forall \theta_i \in [\underline{\theta}_i, \bar{\theta}_i] \forall i = 1, \dots, n$.

Note that the utility function of the agents in this example are given by

$$u_i(f(\theta), \theta_i) = \theta_i y_i(\theta) + t_i(\theta) \forall i = 1, \dots, n.$$

Thus, viewing $y_i(\theta) = v_i(k(\theta))$ will confirm that these utility functions also satisfy the fourth condition required for linear environment. Now we can apply Myerson's characterization theorem to test the Bayesian incentive compatibility of the SCF involved here. It is easy to see that for any bidder i , we have

$$\bar{v}_i(\theta_i) = E_{\theta_{-i}}[v_i(\theta_i, \theta_{-i})]$$

$$\begin{aligned}
 &= E_{\theta_{-i}}[y_i(\theta_i, \theta_{-i})] \\
 &= 1 \cdot P((\theta_{-i})_{(n-1)} \leq \theta_i) + 0 \cdot (1 - P(\theta_i < (\theta_{-i})_{(n-1)})) \\
 &= P((\theta_{-i})_{(n-1)} \leq \theta_i), \tag{20}
 \end{aligned}$$

where $P((\theta_{-i})_{(n-1)} \leq \theta_i)$ is the probability that the given type θ_i of the bidder i is the highest among all the bidders' types. This implies that in the presence of independence assumptions made above, $\bar{v}_i(\theta_i)$ is a non-decreasing function.

We know that for first-price sealed bid auction, $t_i(\theta) = -\theta_i y_i(\theta)$. Therefore, we can claim that for first-price sealed bid auction, we have

$$\bar{t}_i(\theta_i) = -\theta_i \bar{v}_i(\theta_i) \quad \forall \theta_i \in \Theta_i.$$

The above values of $\bar{v}_i(\theta_i)$ and $\bar{t}_i(\theta_i)$ can be used to compute $U_i(\theta_i)$ in following manner.

$$U_i(\theta_i) = \theta_i \bar{v}_i(\theta_i) + \bar{t}_i(\theta_i) = 0 \quad \forall \theta_i \in [\underline{\theta}_i, \bar{\theta}_i]. \tag{21}$$

The above equation can be used to test the second condition of the Myerson's theorem, which require

$$U_i(\theta_i) = U_i(\underline{\theta}_i) + \int_{\underline{\theta}_i}^{\theta_i} \bar{v}_i(s) ds.$$

In view of the Equations (20) and (21), it is easy to see that this second condition of Myerson's characterization theorem is not being met by the SCF used in the first-price sealed bid auction. Therefore, we can finally claim that first-price sealed bid auction is not BIC in linear environment. Q.E.D.

11.5 Second-price sealed bid auction in linear environment

Once again consider the example 3.4 of second-price sealed bid auction. Let us assume that $S_i = \Theta_i = [\underline{\theta}_i, \bar{\theta}_i] \quad \forall i \in N$. In such a case, the second-price auction becomes a direct revelation mechanism $\mathcal{D} = ((\Theta_i)_{i \in N}, f(\cdot))$, where $f(\cdot)$ is an SCF which is same as outcome rule of the second-price auction. We have already seen that this SCF $f(\cdot)$ is DSIC in quasi-linear environment and linear environment is a special case of quasi-linear environment, therefore, it is DSIC in the linear environment also. Moreover, we know that DSIC implies BIC. Therefore, we can directly claim that SCF used in the Vickrey auction is BIC in linear environment.

Table 2 summarizes the properties of the SCFs discussed in above examples in linear environments.

Table 2. Properties of social choice functions in linear environment.

SCF	AE	BB	DSIC	BIC
Fair bilateral trade	✓	✓	×	×
First-price auction	✓	×	×	×
Vickrey auction	✓	×	✓	✓

12. Conclusions

In the first part of this tutorial, we have seen important basic notions and key foundational results in mechanism design. In the second part of this tutorial (Garg *et al* 2008), we will build upon these concepts and results to discuss deeper issues and results in mechanism design theory. The topics that we cover in Part 2 include: (1) Revenue equivalence of auctions; (2) Individual rationality; (3) Moulin mechanisms; (4) Optimal auctions; (5) Characterization of dominant strategy incentive compatible (DSIC) mechanisms; (6) DSIC implementation of Bayesian incentive compatible (BIC) rules; (7) Implementation in ex-post Nash equilibrium; (8) Mechanisms with interdependent types; (9) Implementation of mechanisms; and (10) Other advanced topics in mechanism design.

12.1 To probe further

For a more detailed treatment of mechanism design, the readers are requested to refer to textbooks, such as the ones by Mas-Colell *et al* (1995), Green and Laffont (1979), and Laffont (1988). There is an excellent recent survey article by Nisan (2007). There are many other scholarly survey papers on mechanism design — for example by Myerson (1989) and by Jackson (2001, 2003). The Nobel Prize website has a highly readable technical summary of mechanism design theory (The Nobel Foundation 2007). The recent edited volume on *Algorithmic Game Theory* by Nisan *et al* (2007) also has valuable articles related to mechanism design.

The current paper is not to be treated as a survey on auctions in general. There are popular books (for example, by Milgrom (2004) and Krishna (2002)) and surveys on auctions (for example, (McAfee & McMillan 1987, Milgrom 1989, Paul Klemperer 2004, Wolfstetter 1996, Jayant & Parkes 2005)) which deal with auctions in a comprehensive way.

The current paper is also not to be treated as a survey on combinatorial auctions (currently an active area of research). Exclusive surveys on combinatorial auctions include the articles by de Vries and Vohra (2003, 2005), Pekec and Rothkopf (2003), and Narahari & Dayama (2005). Cramton (2005) has brought out a comprehensive edited volume containing expository and survey articles on varied aspects of combinatorial auctions.

For a more comprehensive treatment of mechanism design and its applications in network economics, the readers are referred to the forthcoming monograph by Narahari *et al* (2008).

We thank Professor Vivek Borkar, Editor, Sadhana, for encouraging us to write this paper. Our grateful thanks to three anonymous referees for their insightful comments and helpful suggestions. The second author would like to acknowledge the support he has received from the Homi Bhabha Fellowships Council, Mumbai and from the Office of Naval Research, Washington, DC.

Notation

n	Number of agents
N	Set of agents: $\{1, 2, \dots, n\}$
Θ_i	Type set of agent i
Θ	Set of all type profiles = $(\Theta_1 \times \dots \times \Theta_n)$
Θ_{-i}	Set of all profiles of types of agents other than $i = (\Theta_1 \times \dots \times \Theta_{i-1} \times \Theta_{i+1} \times \dots \times \Theta_n)$

θ_i	Actual type of agent i , $\theta_i \in \Theta_i$
θ	A profile of actual types $= (\theta_1, \dots, \theta_n) \in \Theta$
θ_{-i}	A profile of actual types of agents other than $i = (\theta_1, \dots, \theta_{i-1}, \theta_{i+1}, \dots, \theta_n) \in \Theta_{-i}$
$\hat{\theta}_i$	Reported type of agent i , $\hat{\theta}_i \in \Theta_i$
$\hat{\theta}$	A profile of reported types $= (\hat{\theta}_1, \dots, \hat{\theta}_n) \in \Theta$
$\hat{\theta}_{-i}$	A profile of reported types of agents other than $i = (\hat{\theta}_1, \dots, \hat{\theta}_{i-1}, \hat{\theta}_{i+1}, \dots, \hat{\theta}_n) \in \Theta_{-i}$
$\Phi_i(\cdot)$	A cumulative distribution function (CDF) on Θ_i
$\phi_i(\cdot)$	A probability density (or mass) function (PDF) on Θ_i
$\Phi(\cdot)$	A cumulative distribution function on Θ
$\phi(\cdot)$	A probability density (or mass) function on Θ
X	Outcome set
x	A particular outcome, $x \in X$
$u_i(\cdot)$	Utility function of agent i
$f(\cdot)$	A social choice function
F	Set of social choice functions
$W(\cdot)$	A social welfare function
\mathcal{M}	An indirect mechanism
\mathcal{D}	A direct revelation mechanism
$g(\cdot)$	Outcome rule of an indirect mechanism
S_i	Set of actions available to agent i in an indirect mechanism
S	Set of all action profiles $= S_1 \times \dots \times S_n$
S_{-i}	Set of all profiles of actions of agents other than agent i
b_i	A bid of agent i ; $b_i \in S_i$
b	A profile of bids $= (b_1, \dots, b_n) \in S$
b_{-i}	A profile of bids by agents other than $i = (b_1, \dots, b_{i-1}, b_{i+1}, \dots, b_n)$
$b^{(k)}$	k^{th} highest element in (b_1, \dots, b_n)
$(b_{-i})^{(k)}$	k^{th} highest element in $(b_1, \dots, b_{i-1}, b_{i+1}, \dots, b_n)$
$s_i(\cdot)$	A strategy of agent i
$s(\cdot)$	A profile of strategies $= (s_1(\cdot), \dots, s_n(\cdot))$
K	A Set of project choices
k	A particular project choice, $k \in K$
t_i	Monetary transfer to agent i
$v_i(\cdot)$	Valuation function of agent i
$U_i(\cdot)$	Expected utility function of agent i
X_f	Set of feasible outcomes
\mathbb{R}	Set of real numbers

References

- Aggarwal G, Goel A, Motwani R 2006 Truthful auctions for pricing search keywords. In *EC'06: Proceedings of the 7th ACM Conference on Electronic Commerce*, (New York, USA: ACM) 1–7
- Altman A, Tennenholtz M 2005 Ranking systems: The pagerank axioms. In *6th ACM Conference on Electronic Commerce (EC'05)* 1–8
- Altman A, Tennenholtz M 2006 Quantifying incentive compatibility of ranking systems. In *Proceedings of The Twenty-First National Conference on Artificial Intelligence* 21: 586–591

- Anderegg L, Eidenbenz 2003 Ad hoc-VCG: A truthful and cost-efficient routing protocol for mobile ad hoc networks with selfish agents. In *9th ACM Annual International Conference on Mobile Computing and Networking (MobiCom'03)*, San Diego, California, USA, September 14–19, 245–259
- Arrow K 1979 The property rights doctrine and demand revelation under incomplete information. In M Boskin, (ed), *Economics and Human Welfare*. (New York: Academic Press)
- Baar T, Olsder G J 1999 *Dynamic Noncooperative Game Theory*. SIAM Series in Classics in Applied Mathematics
- Brandt F 2003 Social choice and preference protection — towards fully private mechanism design. In *4th ACM Conference on Electronic Commerce (EC'03)* 220–221
- Buyya R 2002 *Economic-based distributed resource management and scheduling for grid computing*. PhD thesis, School of Computer Science and Software Engineering, Monash University, Australia
- Caplice C, Sheffi Y 2005 Combinatorial auctions for truckload transportation. In P Cramton, Y Shoham, R Steinberg, (eds.), *Combinatorial Auctions*, (Cambridge, Massachusetts, USA: The MIT Press) 539–572
- Chandrashekar T S, Narahari Y, Rosa C H, Kulkarni D, Tew J D, Dayama P 2006 Auction based mechanisms for electronic procurement. *IEEE Transactions on Automation Science and Engineering* 4(3): 297–321
- Clarke E 1971 Multi-part pricing of public goods. *Public Choice* 11: 17–23
- Cramton P 2005 Simultaneous ascending auctions. In P Cramton, Y Shoham, R Steinberg, (eds.), *Combinatorial Auctions*, (Cambridge, Massachusetts, USA: The MIT Press) 99–114
- Das A, Grosu D 2005 Combinatorial auction-based protocols for resource allocation in grids. In *IPDPS'05: Proceedings of the 19th IEEE International Parallel and Distributed Processing Symposium (IPDPS'05) - Workshop 13*, (Washington, DC: USA, IEEE Computer Society) 251-1
- d'Aspremont C, Gérard-Varet L A 1979 Incentives and incomplete information. *J. Public Econ.* 11: 25–45
- de Vries S, Vohra R V 2003 Combinatorial auctions: A survey. *INFORMS J. Computing* 15(1): 284–309
- de Vries S, Vohra R V 2005 Design of combinatorial auctions. In *Handbook of Quantitative Supply Chain Analysis: Modelling in the E-Business Era*. International Series in Operations Research and Management Science, (Norwell: MA, USA: Kluwer Academic Publishers) 247–292
- Edelman B, Ostrovsky M, Schwarz M 2006 Internet advertising and the generalized second price auction: Selling billions of dollars worth of keywords. In *2nd Workshop on Sponsored Search Auctions in conjunction with the ACM Conference on Electronic Commerce (EC'06)*, Ann Arbor, MI
- Eidenbenz S, Santi P, Resta G 2005 COMMIT: A sender centric truthful and energy-efficient routing protocol for ad hoc networks. In *Workshop on Wireless, Mobile, and Ad hoc Networks (WMAN) in conjunction with 19th IEEE International Parallel and Distributed Processing Symposium (IPDPS'05)*
- Eso M, Kalagnanam J, Ladanyi L, Li Y G 2001 Winner determination in bandwidth exchanges. Technical report, IBM TJ Watson Research Center
- Feigenbaum J and Shenker S 2002 Distributed algorithmic mechanism design: Recent results and future directions. In *6th International Workshop on Discrete Algorithms and Methods for Mobile Computing and Communications (MobiCom'02)*, New York: 1–13
- Feigenbaum J, Papadimitriou C H, Sami R, Shenker S 2002 A BGP-based mechanism for lowest-cost routing. In *21st ACM Symposium on Principles of Distributed Computing (PODC'02)*, New York: June 30–July 4, 173–182
- Garg D 2006 *Design of Innovative Mechanisms for Contemporary Game Theoretic Problems in Electronic Commerce*. PhD thesis, Department of Computer Science and Automation, Indian Institute of Science, Bangalore, India
- Garg D, Narahari Y 2005 Design of incentive compatible mechanisms for stackelberg problems. In *First Workshop on Network and Internet Economics, WINE-2005*, Springer, Lecture Notes in Computer Science, Decemeber 3828: 718–727

- Garg D, Narahari Y, Gujar S 2008 Foundations of mechanism design: A tutorial - Part 2: Advanced concepts results. *Sādhanā* 33: 131–174
- Gautam Raghav, Hemachandra N, Narahari Y, Hastagiri Prakash 2007 Optimal auctions for multi-unit procurement with volume discount bidders. In *Proceedings of IEEE Conference on Electronic Commerce (IEEE CEC-2007)* 21–28
- Gibbard A 1973 Manipulation of voting schemes. *Econometrica* 41: 587–601
- Green J R, Laffont J J 1979 *Incentives in Public Decision Making*. (Amsterdam: North-Holland Publishing Company)
- Grosu D, Chronopoulos A T 2003 A load balancing mechanism with verification. In *IPDPS'03: Proceedings of the 17th International Symposium on Parallel and Distributed Processing*, (Washington, DC: USA, IEEE Computer Society) 163–170
- Grosu D, Chronopoulos A T 2004 Algorithmic mechanism design for load balancing in distributed systems. *Systems, Man, and Cybernetics, Part B, IEEE Transactions on* 34(1): 77–84
- Groves T 1973 Incentives in teams. *Econometrica* 41: 617–631
- Gyongyi Z, Garcia-Molina H 2005 Link spam alliances. In *31st International Conference on Very Large Data Bases (VLDB'05)*, Trondheim, Norway
- Hershberger J, Suri S 2001 Vickrey prices and shortest paths: What is an edge worth? In *42nd Annual IEEE Symposium on Foundations of Computer Science (FOCS'01)* 252–259
- Hurwicz L 1960 Optimality and informational efficiency in resource allocation processes. In *Mathematical Methods in the Social Sciences*. Arrow, Karlin and Suppes (eds.). Stanford University Press
- Hurwicz L 1972 On informationally decentralized systems. In *Decision and Organization*. Radner and McGuire. North-Holland, Amsterdam
- Jackson M O 2001 A crash course in implementation theory. *Social Choice and Welfare* 18: 655–708
- Jackson M O 2003 Mechanism theory. In Ulrich Derigs, (ed.), *Optimizations and Operations Research*. Oxford, UK
- Jayant K, Parkes D C 2005 Auctions, bidding, and exchange design. In David Simchi-Levi, David S Wu, and Max Shen, (eds.), *Handbook of Quantitative Supply Chain Analysis: Modelling in the E-Business Era*, Int. Series in Operations Research and Management Science. (Norwell, USA: Kluwer Academic Publishers)
- John C Harsanyi 1967 Games with incomplete information played by Bayesian players. Part I: The basic model. *Management Sci.* 14: 159–182
- John C Harsanyi 1968a Games with incomplete information played by Bayesian players. Part II: Bayesian equilibrium points. *Management Sci.* 14: 320–334
- John C Harsanyi 1968b Games with incomplete information played by Bayesian players. Part III: The basic probability distribution of the game. *Management Sci.* 14: 486–502
- Kameshwaran S, Narahari Y, Rosa C H, Kulkarni D, Tew J D 2006 Multi-attribute electronic procurement using goal programming. *European J. Operational Res.* 179: 518–536
- Karabuk S, David Wu S 2005 Incentive schemes for semiconductor capacity allocation: A game theoretic analysis. *Production and Operations Management* 14(2): 175–188
- Krishna V 2002 *Auction Theory*. (Burlington, USA: Academic Press)
- Laffont J J 1988 *Fundamentals of Public Economics* (Cambridge: The MIT Press)
- Martin J Osborne 2003 *An Introduction to Game Theory*. (Cambridge: The MIT Press)
- Maskin E 1999 Nash equilibrium and welfare optimality. *Rev. Econ. Studies* 66: 23–38
- Mas-Colell A, Whinston M D, Green J R 1995 *Microeconomic Theory*. (New York: Oxford University Press)
- McAfee P R, McMillan J 1987 Auctions and bidding. *J. Econ. Literat.* 25(2): 699–738
- Milgrom P 1989 Auctions and bidding: A primer. *J. Econ. Perspectives* 3(3): 3–22
- Milgrom P 2004 *Putting Auction Theory to Work*. (New York, USA: Cambridge University Press)
- Myerson R B 1981 Optimal auction design. *Mathematics of Operations Research* 6(1): 58–73
- Myerson R 1989 Mechanism design. In J Eatwell, M Milgate, P Newman, (eds.), *The New Palgrave: Allocation, Information, and Markets*, Norton, New York 191–206
- Myerson R B 1997 *Game Theory: Analysis of Conflict*. (Cambridge, Massachusetts: Harvard University Press)

- Naor M 2001 Cryptography and mechanism design. In *8th Conference on Theoretical Aspects of Rationality and Knowledge (TARK)*, Certosa di Pontignano, University of Siena, Italy
- Naor M, Pinkas B, Sumner R 1999 Privacy preserving auctions and mechanism design. In *EC'99: Proceedings of the 1st ACM conference on Electronic commerce*, (New York, USA: ACM Press) 129–139
- Narahari Y, Dayama P 2005 Combinatorial auctions for electronic business. *Sādhanā* 30(2–3): 179–212
- Narahari Y, Srivastava N K 2007 Incentive compatible mechanisms for decentralized supply chain formation. In *Proceedings of IEEE Conference on Electronic Commerce (IEEE CEC-2007)* 315–322
- Narahari Y, Garg D, Rama Suri N, Prakash H 2008 *Game Theoretic Problems in Network Economics and Mechanism Design Solutions*. Advanced Information and Knowledge Processing Series, (London: Springer)
- Nisan N 2007 Introduction to mechanism design (for computer scientists). In E Tardos, N Nisan, T Roughgarden, V Vazerani, (eds.), *Algorithmic Game Theory* (New York, USA: Cambridge University Press) 209–242
- Nisan N, Ronen A 2000 Computationally feasible VCG mechanisms. In *2nd ACM Conference on Electronic Commerce (EC'00)*, New York 242–252
- Nisan N, Ronen A 2001 Algorithmic mechanism design. *Games and Economic Behaviour* 35: 166–196
- Nisan N, Roughgarden T, Tardos E, Vazerani V 2007 *Algorithmic Game Theory*. Cambridge University Press
- Nisan N 1999 Algorithms for selfish agents, mechanism design for distributed computation. In *16th Symposium on Theoretical Aspects of Computer Science (STACS'99)*, Heidelberg 1–15
- Paul Klemperer 2004 *Auctions: Theory and Practice. The Toulouse Lectures in Economics*. Princeton University Press
- Pekec A, Rothkopf M H 2003 Combinatorial auction design. *Management Sci.* 49: 1485–1503
- Prakash H V, Narahari Y 2005 A strategy-proof auction mechanism for grid scheduling with selfish entities. In *2nd International Conference on Web Information Systems and Technologies (WEBIST'05)*, Satubal, Portugal 178–183
- Roughgarden T 2005 *Selfish Routing and the Price of Anarchy*. (Cambridge: The MIT Press)
- Satterthwaite M A 1975 Strategy-proofness and arrow's conditions: Existence and correspondence theorem for voting procedure and social welfare functions. *J. Econ. Theory* 10: 187–217
- Suri Rama N 2006 Design of incentive compatible broadcast protocols for wireless ad-hoc networks. Technical report, Master's Dissertation, Department of Computer Science and Automation, Indian Institute of Science, Bangalore, India
- The Nobel Foundation 2007 The Sveriges Riksbank Prize in Economic Sciences in memory of Alfred Nobel 2007: Scientific Background. Technical report, The Nobel Foundation, Stockholm, Sweden
- Varian H R 1995 Economic mechanism design for computerized agents. In *1st USENIX Workshop on Electronic Commerce* Berkeley, CA, USA
- Vickrey W 1961 Counterspeculation, auctions, and competitive sealed tenders. *J. Finance* 16(1): 8–37
- Walsh W E, Wellman M P, Wurman P R, MacKie-Mason J K 1998 Some economics of market-based distributed scheduling. In *ICDCS'98: Proceedings of the The 18th International Conference on Distributed Computing Systems*, (Washington, DC, USA: IEEE Computer Society) 612
- Walsh W E, Wellman M P, Ygge F 2000 Combinatorial auctions for supply chain formation. In *2nd ACM Conference on Electronic Commerce (EC'00)*, Minneapolis, Minnesota 260–269
- Walsh W E, Wellman M P 2003 Decentralized supply chain formation: A market protocol and competitive equilibrium analysis. *J. Artificial Intell. Res.* 19: 513–567
- Wang W, Xiang-Yang Li, Wang Y 2004 Truthful multicast routing in selfish wireless networks. In *MobiCom'04: Proceedings of the 10th annual international conference on Mobile computing and networking*, (New York, USA: ACM) 245–259
- Wolfstetter E 1996 Auctions: An introduction. *Econ. Surv.* 10: 367–421