Mechanism Design

1 Introduction

The fundamental problem in economics is how to allocate scarce resources. Many different allocation systems or mechanisms could be used. A very simple allocation system is dictatorship – a single individual (or perhaps a committee of individuals) determines the resource allocation for all individuals. Another simple allocation system is that of "first come, first serve" which simply means that the first person to arrive gets the resource. Typically, economists like to consider how markets or prices allocate resources. The "first come, first serve" allocation system might be coupled with a payment system, like campus parking. The parking fee allows one to park on campus, but does not guarantee the consumer a particular parking space. Many posted-price markets have an element of "first come, first serve" present as well – if there is only one item remaining on the shelf (if we consider a brick and mortar location) or one item remaining in an online store (if we consider virtual locations), then the first person willing to pay the price for that item receives the item. Negotiations, or bargaining, are also systems by which resources are allocated.

While those are all interesting allocation systems, we will focus on a particular type of allocation system, which we will call a "mechanism." Consider a game of incomplete information, where one party (the principal or seller) in a transaction does not know some information (perhaps the willingness to pay) the other party (the agent or buyer) has. Mechanism design can generally be thought of in three steps:

1. A seller designs a mechanism/contract/incentive scheme in which buyers send "messages" and allocations are made based upon the messages sent. An "allocation" here refers not only to the actual good(s) being transferred, but also any transfer payments that are made between buyer(s) and seller. For now our focus will be on simultaneous messages.

2. Buyers choose whether or not they want to participate in the game, or, alternatively one could think about this step as the buyers accepting or rejecting the seller’s proposed mechanism. If buyers choose not to participate in the game they would typically have some reservation level of utility. This step requires consideration of the buyer’s participation constraint (PC) when the buyer determines whether or not to participate. There are some instances in which the "seller" can force the "buyer" to participate, thus removing the need for a participation constraint. Consider the government imposing a tax scheme on individuals or corporations. These economic agents must participate, even if "participating" means "not filing taxes properly and suffering the consequences." That scenario is different than eBay, which has no authority to force individuals to bid in its auctions or punish them for failing to bid in its auctions.

3. The final step is that buyers who choose to participate in the game submit their messages, and then an allocation is made based upon the submitted messages and the mechanism design.

Another key constraint that the seller must consider when designing the mechanism is the buyer’s incentive compatibility (IC) constraint. Sellers want buyers to behave in a certain manner, and a properly designed mechanism can elicit certain types of responses from buyers. Alternatively, in a principal-agent framework, the principal would like to structure a contract to align the incentives of the agent with those of the principal.

\[^1\]It may be helpful to have a particular example in mind. If so, consider the auctions we have discussed, as we will discuss them in more detail throughout this section of notes.

\[^2\]Some texts refer to this constraint as the individual rationality (IR) constraint.
As a specific example of a mechanism consider the first-price sealed bid auction we discussed. The allocation rule is that the bidder who submits the highest bid receives the item and pays an amount equal to his or her submitted bid, while all other bidders make no payment. The "messages" in this mechanism are the bids submitted by the bidders. All messages are used in determining the final allocation and payment because they must all be compared to each other to determine which is the highest bid.

2 Two bidder, two type auction example

Consider a seller with a single good for sale and two bidders who are ex ante identical. Bidder values can be one of two possibilities: \( \tilde{\theta} \) with probability \( p \) and \( \bar{\theta} \) with probability \( \bar{p} \). Let \( s_1 \) and \( s_2 \) be the realizations of the bidder strategies., \( X_i(s_1, s_2) \) be the probability the good is transferred to bidder \( i \), and \( T_i(s_1, s_2) \) be the transfer payment of bidder \( i \) to the seller. Comparing the 1st-price and 2nd-price sealed bid auctions:

- 1st-price: \( X_i(s_1, s_2) = 1 \) if \( s_i > s_j \)
  - 2nd-price: \( T_j(s_1, s_2) = 0 \) if \( s_i > s_j \)

We will restrict our discussion to truth telling mechanisms, the reason for which will be discussed shortly. A truth telling mechanism means that the bidders will reveal their true type (as in the equilibrium of a 2nd-price sealed bid auction). Let \( X \) be the probability of receiving the good and \( T \) be the expected payment when the value is \( \tilde{\theta} \). Let \( X \) be the probability of receiving the good and \( T \) be the expected payment when the value is \( \bar{\theta} \). The seller will structure the mechanism to determine those four parameters based on the submitted signals. We have the following constraints:

- Participation constraints: \( PC_1 : \theta X - T \geq 0 \)
- \( PC_2 : \bar{\theta} X - T \geq 0 \)

- Incentive compatibility constraints: \( IC_1 : \theta X - T \geq \theta X - \bar{T} \)
- \( IC_2 : \bar{\theta} X - T \geq \bar{\theta} X - T \)

Consider what each of these constraints implies. The participation constraints mean that the expected payoff for each type must be at least zero, otherwise they would not participate. The incentive compatibility constraints mean that each type (represented by \( \theta \) and \( \bar{\theta} \)) must be better off submitting a message that reveals his or her type. If \( PC_1 \) are \( IC_2 \) are satisfied, then:

\[
\frac{\theta X}{\bar{\theta} X - T} \geq \frac{T}{\bar{T}}
\]

The next to last step follows because \( X \geq 0 \) and \( \bar{\theta} > \theta \). Thus, \( PC_2 \) is also satisfied and will not be binding unless the seller does not sell to the low type. Now we will show that \( PC_1 \) and \( IC_2 \) are binding constraints. Recall that if the constraints are binding that they can be set to equalities (like the budget constraint in a standard consumer optimization problem). If \( PC_1 \) is not binding, then the seller could increase \( \bar{T} \) and \( T \) by the same amount, make more money, and not violate any constraints. So \( \theta X = \bar{T} \) and a low value bidder does not expect to earn any surplus in equilibrium. Now consider \( IC_2 \). If \( IC_2 \) were not binding then the seller could increase \( T \), earn more money, and not violate any constraints. We can now determine

\[3\] Straightforward adaption from Fudenberg and Tirole, pgs. 246-253. Also see Wolfstetter (1999), pgs. 218-221.
the expected payments from \( PC_1 \) and \( IC_2 \):

\[
\begin{align*}
\theta X &= T \\
\bar{\theta} X - T &= \bar{\theta} X - T \\
\bar{\theta} X - \theta X &= \bar{\theta} X - \theta X \\
\bar{\theta} (X - X) + \theta X &= T \\
\end{align*}
\]

So we have found the payments made, at least in terms of the probabilities. Now we need to find the probabilities.

Now consider the seller’s problem. Letting \( E_0 \) be the seller’s expected utility from the mechanism, then:

\[
E_0 = \rho T + p T
\]

Note that \( E_0 \) is just the payment the seller expects to receive based upon the exogenous probabilities of \( \theta \) and \( \bar{\theta} \) and the expected payments. Substituting for the payments:

\[
\begin{align*}
E_0 &= \rho \theta X + \rho \bar{\theta} (X - X) + \theta X \\
E_0 &= (1 - \rho) \theta X + \rho \bar{\theta} (X - X) + \theta X \\
E_0 &= \theta X + \rho \bar{\theta} (X - X) \\
E_0 &= (\theta - \rho \bar{\theta}) X + \rho \bar{\theta} X \\
\end{align*}
\]

The last part of the problem involves putting constraints on the bidders’ probabilities \( X \) and \( \bar{X} \). Note that if one bidder receives the good then the other bidder does not, so ex ante:

\[
\rho X + \rho \bar{X} \leq \frac{1}{2}
\]

This equation should probably be explained a little more. Recall that \( \bar{\rho} \) and \( \rho \) are exogenous, and that \( \bar{\rho} + \rho = 1 \). Consider that \( \bar{\rho} = 1 \), so that both bidders have value \( \bar{\theta} \). If that is the case, then, if the item is awarded when \( \bar{\theta} \) is revealed, we have \( \bar{X} \leq \frac{1}{2} \). Now, \( \bar{\theta} \) will always be revealed, and there are two bidders, so \( X = \frac{1}{2} \). A similar argument could be made for \( \rho \). Now assume \( \bar{\rho} = \rho = \frac{1}{2} \). We then have \( \frac{1}{2} \bar{X} + \frac{1}{2} \bar{X} \leq \frac{1}{2} \), or \( X + \bar{X} \leq 1 \). Again, \( X \) and \( \bar{X} \) are probabilities, but note that the seller does not have to award the item to a bidder, so the sum could be less than 1 (but never, of course, more than 1).

However there is more to specify. Considering

\[
E_0 = (\theta - \rho \bar{\theta}) X + \rho \bar{\theta} X
\]

there are two potential cases based upon the exogenous parameters \( \theta \), \( \rho \), and \( \bar{\theta} \). It is possible that \( \theta \leq \rho \bar{\theta} \) or \( \theta > \rho \bar{\theta} \). This relationship will determine the optimal probabilities \( X \) and \( \bar{X} \). If \( \theta \leq \rho \bar{\theta} \) then \( E_0 \) is decreasing in \( X \) and the seller wants to set \( X = 0 \). In that case:

\[
E_0 = \rho \bar{\theta} X
\]

The constraint in this scenario is that if both bidders have type \( \bar{\theta} \) then they each must win with probability \( \frac{1}{2} \). Recall that \( X \) is the probability of receiving the item if the bidder reveals \( \bar{\theta} \). Thus, the bidder will always win if the other bidder reveals the low type or will win one-half of the time when the other bidder reveals the high type. So \( X = \frac{\rho + \bar{\rho}}{2} \) because the seller wants to maximize the payment \( T \). The optimal mechanism in this scenario awards the item to no one if both announce \( \theta \), to the bidder who reveals \( \bar{\theta} \) if one reveals the low type and the other the high type, and to each bidder with probability \( \frac{1}{2} \) if they both reveal \( \bar{\theta} \).

When \( \theta > \rho \bar{\theta} \) then \( E_0 \) is strictly increasing in \( X \) and \( \bar{X} \). Then \( \rho X + \rho \bar{X} = \frac{1}{2} \) and

\[
E_0 = \frac{1}{2 \rho} (\theta - \rho \bar{\theta}) + \frac{\rho}{\rho} (\bar{\theta} - \theta) \bar{X}.
\]
While more complicated than \( E\mu_0 = p\theta X \), the form is the same, so \( X = p + \frac{p}{2} \). Using \( pX + p\frac{X}{2} = \frac{1}{2} \) we can see that \( X = \frac{p}{2} \). In this scenario, if both announce the low type then they each receive the item with probability \( \frac{1}{2} \); if one bidder announces the high type and the other bidder announces the low type then the bidder who announces the high type receives the item, and if both announce the high type then they each receive the item with probability \( \frac{1}{2} \).

3 General Results

Among the many results for Bayesian games there are two that are used quite a bit. The first result is the Revelation Principle and the second result is the Revenue Equivalence Theorem.

3.1 Revelation principle

Suppose that a seller wishes to sell an object using some mechanism – the precise mechanism is left unspecified as long as the following conditions are met:

1. The buyers simultaneously make claims about their types. Buyer \( i \) can claim to be any type from his feasible set of types.
2. Given the buyers' claims, buyer \( i \) pays an amount that is a function of all the reported types and receives the good with some probability based upon the reported types (in an auction the bidder with the highest reported type receives the good with probability 1).

Games that satisfy these criteria are known as direct mechanisms, because the only action is to submit a claim about a type. A 1st-price sealed bid auction and a 2nd-price sealed bid auction are direct mechanisms; the all-pay auction example/experiment we did in class is a direct mechanism; the ascending and descending clock auctions, as well as the oral outcry example we did, are not direct mechanisms. I am going to state a variety of forms of the revelation principle from multiple sources to give you all an idea of how it has been used:

**Proposition 1** (MWG pg. 493) Denote the set of possible states by \( \Theta \). In searching for an optimal contract, the owner can without loss restrict himself to contracts of the following form:

1. After the state \( \theta \) is realized, the manager is required to announce which state has occurred.
2. The contract specifies an outcome \( [w(\hat{\theta}), e(\hat{\theta})] \) for each possible announcement \( \hat{\theta} \in \Theta \). (Note that \( w \) and \( e \) are just wages and efforts.)
3. in every state \( \theta \in \Theta \), the manager finds it optimal to report the state truthfully.

**Proposition 2** (MWG pg. 884) Suppose that there exists a mechanism \( \Gamma = (S_1, ..., S_I, g(\cdot)) \) that implements social choice function \( f(\cdot) \) in Bayesian Nash equilibrium. Then \( f(\cdot) \) is truthfully implementable in Bayesian Nash equilibrium.

**Proposition 3** (Fudenberg and Tirole, pg. 255) The principal can content herself with "direct" mechanisms, in which the message spaces are the type spaces, all agents accept the mechanism in step 2 regardless of their types, and the agents simultaneously and truthfully announce their types in step 3.

**Proposition 4** (Gibbons, 1992, pg. 165) Any Bayesian Nash equilibrium of any Bayesian game can be represented by an incentive-compatible direct mechanism.

**Proposition 5** (Wolfstetter, 1999, pg. 214) For any equilibrium of any auction game, there exists an equivalent incentive-compatible direct auction that leads to the same probabilities of winning and expected payments.
Alternatively, one could consult Myerson (1979). Incentive Compatibility and the Bargaining Problem. *Econometrica* 47, 61-73. You may want to read it just to see the simplest theorem ever. Theorem 2: \( F^{**} = F^* \).

Those are just five statements of the revelation principles, some of which are a little general. I highly suggest reading the text surrounding the statement of the theorems to understand what assumptions are being made (for example, in Wolfstetter, what constitutes an "auction game"). Why is the revelation principle useful and important? It is useful because it might be difficult to determine the equilibrium in one game, but we know that if an equilibrium exists in that game then we can find a direct, truth-telling mechanism that has the same general properties. In the next section results we examine the 1st-price sealed bid auction (direct mechanism, not truth-telling) and the 2nd-price sealed bid auction (direct mechanism, truth-telling) in more detail.

### 3.2 Revenue equivalence

In the previous set of notes I provided a comparison between the expected revenue of a 1st-price sealed bid auction and a 2nd-price sealed bid auction. That comparison was made under a specific set of assumptions. A more general statement about the expected revenue from auctions follows:

**Proposition 6** 23.D.3 (Revenue Equivalence Theorem) Consider an auction setting with \( I \) risk-neutral buyers, in which buyer \( i \)'s valuation is drawn from an interval \( [\theta_i, \theta_i] \) with \( \theta_i \neq \theta_i \) and a strictly positive density \( \phi_i(\cdot) > 0 \), and in which buyer’s types are statistically independent. Suppose that a given pair of Bayesian Nash equilibria of two different auction procedures are such that for every buyer \( i \) : (i) For each possible realization of \((\theta_1, ..., \theta_I)\), Buyer \( i \) has an identical probability of getting the good in the two auctions; and (ii) Buyer \( i \) has the same expected utility level in the two auctions when his valuation for the object is at its lowest possible level. Then these equilibria of the two auctions generate the same expected revenue for the seller.

This result relies upon the previous result (revelation principle). The general idea is that there are direct mechanisms that are not truth-telling mechanisms, but they have the same Bayes-Nash equilibrium as a truth-telling mechanism by the revelation principle. We can then compare the expected revenue from the truth-telling mechanisms because the expected revenue relies on the underlying probability distribution of values.

It is important to consider the assumptions that are embedded in that proposition beyond the two that are explicitly enumerated (a buyer who has a specific type in each mechanism has the same probability of winning in each auction and buyers with the lowest possible value have the same expected utility in each auctions). Buyers are risk-neutral. Values are drawn from the same probability distribution. Value draws are statistically independent. Those assumptions are just the SIPV-RN environment mentioned earlier.

While these assumptions are restrictive, they establish a benchmark for comparing more realistic settings. What happens if buyers are not risk-neutral? What happens if the distribution is not symmetric? What happens if value draws are not independent? What happens if value draws are not private?

### 3.3 Winner’s curse

Suppose that I am auctioning off a jar of coins. The jar is see-through, so that you all can see there are coins in the jar. I tell you they are all U.S. coins from 1965-present (prior to 1965 some U.S. coins, notably dimes and quarters, are made of silver and are worth more than their monetary denomination) and you can see various coins (pennies, nickels, dimes, and quarters) in the jar. However, no one is allowed to look inside the jar or take the coins out of the jar. I conduct a 1st-price sealed bid auction for the jar of coins, where the winner gets the coins. Clearly, the monetary amount that each individual would receive is the same because the coins do not depend on the winning bidder. Bidders may have different utility for the coins because perhaps they do not want to carry around coins to spend, but let us assume that they are all students who will happily take money in coin form. Alternatively, we can consider that the bidders have no disutility from the monetary unit and only care about the value of the money. How are individuals’ values formed for this jar of coins?
This auction is different than the ones we have discussed previously. In the prior auctions bidders had different values for the same item. It is fairly simple to motivate that example - there are plenty of goods for which you and your friends would pay different amounts. However, in the jar of coins example, all bidders have the same value for the same item, but they likely have different estimates of the item’s value. They will not know if their estimate is correct unless they win the item and take possession of it. The jar of coins example seems a bit contrived – after all, who would auction off a jar of coins? But there are plenty of examples that fit this particular type of value determination. Consider a seller who has discovered that there is oil on his property. The seller wants to sell because he does not know much about extracting oil and refining it, but the bidders will not know exactly how much oil is in the deposit until they own the property and can begin to extract it. Or, perhaps more relevant to finance students, consider a target firm that is for sale. Other firms would like to buy this target firm, and they have an estimate of the target firm’s value based on observable information, but they will not truly know the target firm’s value until it is acquired.

Auctions of this type are known as "common value auctions." They are different than "private value auctions" because bidders now have a signal about the items value, but do not know the true value until it is purchased. Making the concept slightly more formal:

1. There is a common value $V$ for the item, which is drawn from some underlying probability distribution with support $(v, v)$.

2. Bidders receive a signal $S_i$ of $V$ prior to bidding. The signal $S_i$ is drawn from some probability distribution $(V - \varepsilon, V + \varepsilon)$. Thus, each bidder’s signal is dependent on the common value $V$, but they all have (potentially) different signals.

We will not dive deeply into the determination of this equilibrium, but will consider an example assuming a 1st-price sealed bid auction. If we assume a symmetric equilibrium, then the bidder with the highest signal will win. However, bidders who win know that they have the highest signal and, the more bidders in the auction or the larger $\varepsilon$ is, the more likely that signal is an overestimate of $V$. Thus, they need to shade their bid not only to receive a surplus (as they would in a private value 1st-price sealed bid auction) but also because they realize their signal, if they win, is likely an overestimate of $V$. In equilibrium, bidders in a common value auction make positive profit. However, bidders who are not familiar with the common value auction may (likely) end up bidding too much for the item, and ultimately lose money because the pay some price $P > V$. Overbidding in a common value auction and losing money is know as the winner’s curse. The phrase "winner’s curse" appears in finance journals and I wanted you all to be familiar with it. The term really applies to common value auctions, though some individuals use it (likely incorrectly) with private value auctions. In a private value auction individuals know their values, and can avoid the winner’s curse by making sure they do not submit bids that would lead to payments greater than their value, whereas in common value auctions bidders will likely be submitting a bid below their signal, but that bid might still be above the common value $V$.

---

4See Wolfstetter (1999), pgs. 226-229 for more information.