Theory Generals Study Group
Review Notes: Contracts I

Arda Gitmez Kosti Takala

February 27, 2015

Contents

1 A Very Brief Introduction to Mechanism Design

For simplicity, or what Bengt would call inability to handle the general case, we mostly play with the setup:

- N={1,2,. . . ,n} (set of players)
- $\Theta := \Theta_1 \times \ldots \times \Theta_n$, M, and X (sets of types, messages, and choices, respectively)
- $F = \prod_{i=1}^{n} F_i$, where $F_i : \Theta_i \mapsto [0, 1] \forall i$ (distribution of types, assumes independence)
- $x : M \mapsto X$ and $t : M \mapsto \mathbb{R}^{\ltimes}$ (choice and (monetary) transfer functions, respectively)
- $U_i(x, t, \theta) = u_i(x_i, \theta_i) + t_i$, where $u_i : X \times \Theta_i \mapsto \mathbb{R}$ (quasi-linear[1] utility functions, private values)

1.1 Revelation Principle

The following little theorem simplifies the mechanism design problem by allowing us to restrict our attention to direct mechanisms.

Theorem (Revelation Principle): Any equilibrium allocation[2] of any mechanism with an arbitrary message space M can be truthfully implemented by a direct mechanism ($M_i = \Theta_i$ and everyone is truthful).

[1] What is important here is the separability of money. However, we allow a general Bernoulli utility function.

[2] When we say 'allocation', we are usually referring to the decision that the choice function spits out based on the players' reports, but it can also include the transfers $t(m)$. However, in efficient mechanism design, the transfers are of minor importance in themselves, as long as the efficient decision is made and the budget is balanced.

Proof: Suppose the mechanism we are considering had an equilibrium $(s_1^*(\theta_1), \dots, s_n^*(\theta_n))$ and a respective allocation $(x_1(s^*(\theta)), \dots, x_n(s^*(\theta)))$. Now, the mechanism designer could design a direct mechanism so that whenever i reports his type θ_i, the designer plays $s_i^*(\theta_i)$ for her. Then $(\theta_1, \dots, \theta_n)$ is clearly an equilibrium of the direct mechanism and it leads to the same allocation. Otherwise there would be a type θ_i of player i who had a profitable deviation θ_i'. But then this type would also have a profitable deviation in the game induced by the original mechanism, a contradiction. □

2 Efficient Mechanisms

When we talk about efficient mechanism design, we are interested in knowing what the optimal decions (from the set X) are and whether or not they are implementable by a direct mechanism (or by any mechanism for that matter, thanks to Revelation Principle). Usually we consider dominant incentive compatible (DIC) or Bayesian incentive compatible (BIC) implementation, where the equilibrium concept is either in dominant or Bayesian strategies.

2.1 Vickrey-Clark-Groves Mechanism

VCG mechanisms are a class of incentive compatible (IC) direct mechanisms that implement the first-best decision rule $x^*(\theta)$.[3]. This is done by setting the transfer rule as follows:

$$t_i(\theta) = \sum_{j \neq i} u_j(x_j(\theta), \theta_j) + h_i(\theta_{-i}) \tag{1}$$

There are two important things to note about this transfer rule. The first term makes i the residual claimant of the full social surplus, whereas the second one doesn't affect i's decision/incentives in any way (since it only depends on the reports of the other players). This is why the first-best allocation rule is implementable. Notice, however, that $\sum_{i=1}^n t_i(\theta) = (n-1)\sum_{i=1}^n u_i(x_i(\theta), \theta_i) + \sum_{i=1}^n h_i(\theta_{-i})$, so we have to construct the functions h_i so as to make $\sum_i t_i(\theta) \leq 0$ and, thereby, not run a deficit. This is generally not possible under DIC, but if we are only considering BIC implementation, there is a way to achieve efficiency and budget balance (BB) using something called AGV mechanism (after d'Aspremont and Gerard-Varet). This is because, under BIC, agents only maximize $\mathbb{E}_{\theta_{-i}}[u_i(x_i^*(\theta_i', \theta_{-i}), \theta_i)] + \bar{t}_i(\theta_i')$, where $\bar{t}_i(\theta_i')$ is the expectation of i's transfer given her observing her true type θ_i and reporting θ_i'. This allows the designer to subtract from i's transfer other players' expectations $\bar{t}_j(\theta_j)$ because these obviously are independent of i's report. Thus, the AGV mechanism sets $t_i(\theta) = \bar{t}_i(\theta_i) - \frac{1}{N-1}\sum_{j \neq i} \bar{t}_j(\theta_j)$ – and $\sum_i t_i(\theta) = 0$.

Things to note:

- DIC ⇒ BIC (if $\theta_i \in BR_i(\theta_{-i}) \forall \theta_{-i}$, then θ_i is a best response to a convex combination of θ_{-i})
- Under regularity conditions (those needed for Payoff Equivalence Theorem below), any transfer rule implementing the first-best is a VCG transfer rule (since, by a corollary of payoff equivalence, the transfer rules can differ only by a constant which can be added to $h_i(\theta_{-i})$)
- Groves' scheme sets $h_i(\theta_{-i}) = 0$ (IC and IR, **not BB**)
- Pivot mechanism sets $h_i(\theta_{-i}) = -\max_x \sum_{j \neq i} u_j(x, \theta_{-i})$ (IC and IR, **not BB**)
- AGV mechanism is IC and BB but **not IR**

[3] $x^*(\theta)$ maximizes $\sum_{i=1}^n u_i(x_i(\theta), \theta_i)$

2.2 Payoff Equivalence

Next we will state and prove an extremely useful result that will be used time and time again. The usefulness of the result comes partly from the fact that it allows us to restrict our attention to finding the optimal $x(.)$, since then the transfer scheme follows by the result. The version we consider here is for $n = 1$ and one-dimensional types, but the result generalizes to many players and multidimensional type spaces.

Payoff Equivalence Theorem (n=1): Suppose that $\Theta = [\underline{\theta}, \overline{\theta}] \subset \mathbb{R}$ and that $u_\theta(x, \theta)$ exists and is uniformly bounded[4] for all (x, θ). If the pair $(x(.), t(.))$ is IC, then for all $\theta \in \Theta$:

$$u(x(\theta), \theta) + t(\theta) = u(x(\underline{\theta}), \underline{\theta}) + t(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} u_\theta(x(s), s)ds. \tag{2}$$

Proof: By IC, the value function is $V(\theta) = u(x(\theta), \theta) + t(\theta) = \max_{\theta'}\{u(x(\theta'), \theta) + t(\theta')\}$. We see that V is B-Lipschitz, and therefore continuous, because of the boundedness of u_θ. This implies that V is also absolutely continuous. Next, for all $\theta, \theta' \in (\underline{\theta}, \overline{\theta})$ we have $u(x(\theta), \theta') - u(x(\theta), \theta) \le V(\theta') - V(\theta)$ (the transfers cancel out on the LHS). This implies (by taking limits from above and below) that $V'(\theta) = u_\theta(x(\theta), \theta)$. The claim follows from this combined with V being absolutely continuous. □

There are two important corollaries of the above result. Assuming that two mechanisms are IC and have the same allocation rule $x(.)$, then:

- the value functions of a given type θ differ only by a constant in the two mechanisms (if for some type, the value functions are the same, then this is true for all types)
- the transfer rules differ only by a constant (which is equal to, for example, the difference in the lowest type's value functions under the two mechanisms).

We need to note that if the type spaces aren't "connected" (leading case being discrete type spaces), the transfer rule need not be unique (as opposed to the one obtained by payoff equivalence).

Payoff Equivalence Theorem for BIC mechanisms: Suppose that $\Theta_i = [\underline{\theta_i}, \overline{\theta_i}] \subset \mathbb{R}$ and that $u^i\theta(x, \theta)$ exists and is uniformly bounded for all (x, θ_i). If the pair $(x(.), t(.))$ is BIC, then for all i and all θ_i:

$$\mathbb{E}\big[u^i(x(\theta), \theta_i) + t^i(\theta)\big] = \mathbb{E}\big[u^i(x(\underline{\theta}_i, \theta_{-i}), \underline{\theta}_i) + t^i(\underline{\theta}_i, \theta_{-i})\big] + \int_{\underline{\theta}_i}^{\theta_i} \mathbb{E}\big[u^i_\theta(x(s, \theta_{-i}), s)\big]ds. \tag{3}$$

The two corollaries from above apply to BIC mechanisms, too. If two mechanisms implement the same decisions rule and are BIC, then their expected transfers differ only by a constant. Now we are ready to consider the famous Myerson-Satterthwaite Theorem.

Myerson-Satterthwaite Theorem: Suppose that there is a seller (1) and a buyer (2) whose valuations for a good are $\theta_1 \sim F_1$ on $[\underline{\theta}_1, \overline{\theta}_1]$ and $\theta_2 \sim F_2$ on $[\underline{\theta}_2, \overline{\theta}_2]$, respectively (assume full support). Suppose also that trade is not always efficient nor always inefficient (i.e. that the two intervals of possible valuations overlap). Then there **does not** exist a mechanism that is efficient, BIC, IR, and BB.

Proof: Let $x_i(\theta)$ be the probability that party i gets the good, and let the interim expected utilities of the two parties be: $U_1(\theta_1) = \bar{t}_1(\theta_1) - \overline{x}_1(\theta_1)\theta_1$ and $U_2(\theta_2) = \overline{x}_2(\theta_2)\theta_1 - \bar{t}_2(\theta_2)$. Consider the Pivot mechanism which is efficient, BIC and IR. By payoff equivalence, any efficient BIC mechanism has an expected transfer $\bar{t}_i(\theta) = \bar{t}_i^P(\theta) + c_i$ for all i and θ_i (where $\bar{t}_i^P$ is the interim expected transfer under Pivot). But the interim expected utilities of types $\overline{\theta}_1$ and $\underline{\theta}_2$ are zero under Pivot and it runs a deficit of $\mathbb{E}[max\{\theta_2 - \theta_1, 0\}] = \int_{\underline{\theta}_1}^{\overline{\theta}_2}(1 - F_2(\theta))F_1(\theta)$, [5]. This is positive, but we can't finance it using transfers because of the IR constraints of the two extreme types. Thus, the claim is true. □

[4]So that $\exists B < \infty$ such that $|u_\theta(x, \theta)| \le B$ for all x and all θ.

[5]This is the general deficit of any efficient, BIC and IR mechanism, taken from Myearson and Satterthwaite (1983).

3 Optimal Mechanisms

The previous section dealt with the case of efficient mechanisms, where the decision rule (the efficient allocation) is known *a priori*. This makes the analysis easier in some sense, as the only question that remains to be answered is: "How can we implement this allocation?"[6] But what happens when the objective is, say, maximizing profit rather than social welfare? Then the optimal allocation rule is less transparent, let alone the question of how to implement it.[7] This requires the usage of some restrictions in these type of problems, in order to impose some regularity on the optimal allocation rule. The standard way to go is to assume ***Single-Crossing*** property, which brings a lot of simplifications to the (potentially very cumbersome) Incentive Compatibility constraints. Once these simplifications are achieved, the solution of the problem is pretty mechanical (and frankly, pretty fun once one knows how to do it.) We'll begin by a toy example to see the intuition and then derive generalizations.

3.1 An Example

This is the standard "procurement problem" due to Mirrlees, in its most simple, stripped-down version. There is a an authority (the buyer, B) which needs procurement of a single type of good from a supplier (the seller, S). The value that B attributes to procurement of x units of good is $V(x)$. The seller has a type θ which is privately known by the supplier. We will use the convention that higher θ correspond to lower costs of producing the same unit of goods (i.e. higher types.) In particular, type θ's cost of producing x units for is $C(x, \theta)$. Thus, when there is a monetary transfer t from B to S, the utilities are given by:

$$u_B(x, t) = U(x) - t \tag{4}$$
$$u_S(x, t, \theta) = t - C(x, \theta) \tag{5}$$

Throughout this example we'll retain the following assumptions:

- $C(0, \theta) = 0$. Just a normalization.
- For any $x > 0$, $\theta_1 < \theta_2$ implies that $C(x, \theta_1) > C(x, \theta_2)$. In other words, high types are more cost-efficient. This is to impose an order on the types.
- $C_x > 0$, $C_{xx} > 0$, $U' > 0$, $U'' \leq 0$. Standard assumptions.
- $C_{x\theta} < 0$. In other words, higher θ does not only mean lower cost, but it also means a *flatter cost curve* everywhere. This is the very crucial *Single-Crossing* property, which is at the core of any application about optimal mechanisms. One thing to realize is that Single-Crossing is a *global* property: it's about the general shape of the cost curves, rather than properties at specific (local) points. The globality of this property allows us to reduce the the global incentive compatibility requirements to the local ones.

A simple remark that one can make at this point is that the efficient (first-best) allocation rule is the one which solves:

$$x^{FB}(\theta) = \arg\max_x \quad U(x) - C(x, \theta)$$

Which leads to the characterization $U'(x^{FB}(\theta)) = C_x(x^{FB}(\theta), \theta)$. Clearly, under the assumptions we make, $x^{FB}(\theta)$ must be increasing in θ.

For this example, we will restrict attention to the two-type case, which conveys most of the intuition and the properties of the solution in the general case. Suppose $\Theta = \{\theta_1, \theta_2\}$, with $\theta_1 < \theta_2$; and let $Prob\{\theta = \theta_1\} = p$,

[6] To which the answer is: "Use Payoff Equivalence Theorem."
[7] Yet the answer is still "Use Payoff Equivalence Theorem."

$Prob\{\theta = \theta_2\} = 1 - p$, where $p \in [0, 1]$. In this case, the buyer is simply trying to maximize the expected payoff $p(U(x_1) - t_1) + (1 - p)(U(x_2) - t_2)$.

So, how to attack this problem? We'll postulate two possible approaches.

Approach 1. (The naive approach.) The basic intuition suggests that the buyer should be offering a transfer schedule $t(x)$ (think of this as a declaration of the form: "I'll pay $t(x)$ to anyone who brings me x." which is probably what will happen in practice.) Once this schedule is offered, the seller (which has information on θ) would then pick its production level to solve:

$$x(\theta) = \arg\max_{x} \quad t(x) - C(x, \theta)$$

The buyer can then infer what choices will be made by each type if $t(x)$ is offered, and solve back for optimal schedule. Note, however, that this tends to be a very complicated problem, because the set of possible price schedules is a very large set (it contains every single function of x!) So, even though in principle this is solvable, in practice this is rarely done. Nevertheless, there are some observations we can make using this version of the problem. First of all, realize that the seller's problem in this form is supermodular in (x, θ), thanks to the Single-Crossing property. By Monotone Selection Theorem, this implies that high types will pick higher x's given any schedule $t(x)$. Within the context of this example, this means that we will necessarily have $x(\theta_1) \leq x(\theta_2)$, i.e. an increasing production scheme in equilibrium. This is indeed a constraint imposed by the assumption of Single-Crossing cost schedules: *any solution of this problem must involve* $x(\theta_1) \leq x(\theta_2)$. One could indeed also prove that the converse also holds: *any non-decreasing production scheme* $x(\theta_1) \leq x(\theta_2)$ *can be implemented by some transfer schedule.* A picture explains it all, demonstrating how (x_1, x_2) can be implemented by the proper choice of transfers (t_1, t_2):

The narrative in the previous paragraph, which reduces to the suggestion of setting a separate price-quantity pair (x_i, t_i) for each type, brings us to the second approach.

Approach 2. (The "mechanism designer" approach.) This seemingly small, and somewhat obvious, step took years to discover, and is the single most important significant contribution of the theory of mechanism design. Instead of offering a whole transfer scheme, the buyer can simply offer a menu $\{(x(\theta_1), t(\theta_1)), (x(\theta_2), t(\theta_2))\}$ (think of this as a declaration of the form: "Just tell me your type and I'll tell you what to produce, and what you'll get as a compensation."[8] Probably, this is not an accurate description of what happens in practice, but who cares? We get a solution! Plus, we can always use the Taxation Principle to find the transfer scheme which implements the same outcome, once the solution is found.) For notational simplicity, we'll use (x_i, t_i) to denote $(x(\theta_1), t(\theta_1))$ and $C_i(x)$ to denote $C(x, \theta_i)$. Also, we'll assume that the firms have the outside option of producing nothing and receiving no transfer.[9]

[8] Realize the revelation principle at work here!

[9] We're also assuming that the firms make their decision to accept the contract *after* their types are realized. If the firm draws her type after the contract is accepted but before production takes place, then we would have an ex ante IR constraint (the one which requires the expected profit to be nonzero) instead of the interim IR constraint (the one that requires the profit

In other words, the real contribution of mechanism design theory is to point out that the problem can be written in the (more tractable) form:

$$\begin{aligned}
\max_{(x_1,t_1),(x_2,t_2)} \quad & p(U(x_1)-t_1)+(1-p)(U(x_2)-t_2) & \\
\text{s.t.} \quad & t_1 - C_1(x_1) \geq t_2 - C_1(x_2) & (IC_1)\\
& t_2 - C_2(x_2) \geq t_1 - C_2(x_1) & (IC_2)\\
& t_1 - C_1(x_1) \geq 0 & (IR_1)\\
& t_2 - C_2(x_2) \geq 0 & (IR_2)
\end{aligned}$$

This is a pretty easy problem which can be solved using standard arguments.[10] In particular,

- Realize that (IC_1) and (IC_2) imply: $C_1(x_2) - C_1(x_1) \geq C_2(x_2) - C_1(x_1)$, which, along with Single-Crossing, implies that we need to have $x_2 \geq x_1$.
- One can show that (IC_2) and (IR_1) imply (IR_2) already, so that (IR_2) is slack. To see this, just write: $t_2 - C_2(x_2) \geq t_1 - C_2(x_1) \geq t_1 - C_1(x_1) \geq 0$, where the first inequality follows from (IC_2), the second follows from the assumptions on cost structure, and the third follows from (IR_1). So we can just omit (IR_2) as long as we keep (IR_1) and (IC_2).
- Given that (IR_2) is slack, a simple variational argument shows that (IC_2) must bind (Otherwise decrease t_2. This keeps (IC_2) satisfied, relaxes (IC_1) and doesn't affect (IR_1), while increasing the value of objective function.) This implies that $t_2 = C_2(x_2) + t_1 - C_2(x_1)$, pinning down one of the variables.
- Using the value of t_2 derived in the previous step, and that $x_2 \geq x_1$, one can show that (IC_1) is slack. Just substitute the equality into (IC_1) and see that it holds under Single-Crossing. Thus we can ignore (IC_1) as long as $x_2 \geq x_1$.
- Finally, realize that (IR_1) must bind. This is because otherwise one can decrease t_1 and improve the objective function value. This implies that we must have $t_1 = C_1(x_1)$, pinning down another variable.

All in all, the problem is simplified into the following:

$$\begin{aligned}
\max_{(x_1,t_1),(x_2,t_2)} \quad & p(U(x_1)-t_1)+(1-p)(U(x_2)-t_2) & \\
\text{s.t.} \quad & t_2 = C_2(x_2) + t_1 - C_2(x_1) & (IC_2)\\
& t_1 = C_1(x_1) & (IR_1)\\
& x_2 \geq x_1 & (MON)
\end{aligned}$$

Which, after substitutions, becomes:

$$\begin{aligned}
\max_{x_1,x_2} \quad & p[U(x_1) - C_1(x_1) - \frac{1-p}{p}(C_1(x_1) - C_2(x_1))] + (1-p)[U(x_2) - C_2(x_2)] & \\
\text{s.t.} \quad & x_2 \geq x_1 & (MON)
\end{aligned}$$

The approach that is followed after one obtains this problem is well-known to anyone who has solved any adverse selection problem at some point: ignore (MON), solve the relaxed problem, and cross your fingers hoping that the solution to the relaxed problem satisfies (MON). Following this route, we end up with first-order conditions characterizing the optimal quantities x_1^* and x_2^*:

to be nonzero for each type). This would clearly change the nature of optimal contract: See 14.281, Fall 2014, PSet 6 Question 1. Even though that's a question on moral hazard, the ideas go through.

[10] The way in which these arguments are presented (even the order in which the constraints are eliminated) is also pretty standard. Stole's lecture notes present the two-type case for the monopoly pricing model (p.49), and you can see the paralellity.

$$U'(x_2^*) = C_2'(x_2^*)$$
$$U'(x_1^*) = C_1'(x_1^*) + \frac{1-p}{p}(C_1'(x_1^*) - C_2'(x_1^*))$$

One can compare these equations with the one characterizing the first-best levels ($U'(x_i^{FB}) = C_i'(x_i^{FB})$) to realize that: $x_2^* = x_2^{FB}$ and $x_1^* < x_1^{FB}$. But we already argued that $x_1^{FB} < x_2^{FB}$, thus $x_1^* < x_2^*$ and (MON) is satisfied at the relaxed problem's solution, so this is indeed a solution! This means that we've fully solved the problem (t_1^* and t_2^* can be calculated using (IR_1) and (IC_2)).

We will see that most of the ideas presented in this example and many properties of the solution are fairly generalizable. The recurrent themes are:

- Thanks to the Single-Crossing property, the general problem can be relaxed such that high type's IC constraint binds, and low type's IR constraint binds. The simplification adds a MON constraint to the problem, but it is almost always ignored and hoped that the solution of the relaxed problem satisfies MON.
- In the end, the relaxed problem looks like a combination of several first-best problems. In particular, ignoring MON, one can write the above problem as two separate problems (or, as the contract theorists call it, *solve pointwise*):

$$\max_{x_1} \quad U(x_1) - C_1(x_1) - \frac{1-p}{p}(C_1(x_1) - C_2(x_1))$$

  and

$$\max_{x_2} \quad U(x_2) - C_2(x_2)$$

  These two suspiciously look like problems to solve for first-best allocations, only with somewhat modified cost functions (the high type's cost function remains the same, and the low type has an additive cost term, coming from the IC constraint of the high type.) Intuitively, one can think of these additional terms as the costs that arise from the information asymmetry: there is an additional cost of low type producing the good, because the high type has the ability to imitate the low type (thus we need to pay a cost to keep the high type from imitating the low type). These modified cost functions are what Myerson calls as *virtual cost*: the costs that are inclusive of the information rents. The introduction of virtual costs allow us to solve the problem pointwise, which is the main methodological contribution of Myerson.
- The (newly-introduced) intuition of virtual costs indeed helps us to come up with new insights and comparatice statics. For instance, realize that a lower p (i.e. a higher probability of high type) increases the virtual cost of the low type. Indeed, for sufficiently low p, it may be optimal to set $x_1^* = 0$ altogether! Intuitively this is equivalent to saying "There are so many people who want to imitate you that asking you to produce something is too costly. You better leave the market and leave me alone with those guys."
- The optimal solution involves: $x_2^* = x_2^{FB}$ and $x_1^* < x_1^{FB}$. The high type produces efficient amount and the low type produces less-than-efficient. *There are downward distortions.*
- The low type makes zero profits at the optimal solution ($t_1^* = C_1(x_1^*)$) and the high type receives positive profits ($t_2^* > C_2(x_2^*)$). *High type receives an information rent.* This is because one needs to leave rents to high type to keep her from imitating low type.
- All in all, the main trade-off is ***rents vs productive efficiency***. This is the general theme of adverse selection problems, just like *risk vs incentives* is the theme of moral hazard problems.

As one can observe, the introduction of Single-Crossing property is the one thing that really helps us here. For the more general problems, a better comprehension of this property will definitely help. Such armament is built in the next section.

3.2 Single-Crossing and Constraint Simplification

I'll just straightforwardly present what's needed for the practical use: the simplification theorem that allows us to go from global constraints to local ones under Single-Crossing condition.

Let $x \in X$ be the decision, $\theta \in \Theta$ to be the agent's type, and $t \in \mathbb{R}$ to be the transfer from the agent to the principal. The agent's utility is:

$$u(x, t, \theta) = v(x, \theta) + t \tag{6}$$

We are assuming that X and Θ are totally ordered sets. Having X,Θ as subsets of $\mathbb{R}$ will suffice; indeed, we'll assume that $\Theta = [\underline{\theta}, \bar{\theta}]$.

Definition 1. *The function $v : X \times \Theta \to \mathbb{R}$ has **strictly increasing differences (SID)**[11] if, for each $x_1 < x_2$, $\theta_1 < \theta_2$,*

$$v(x_2, \theta_2) - v(x_1, \theta_2) > v(x_2, \theta_1) - v(x_1, \theta_1) \tag{7}$$

Intuitively, this is the requirement that the payoff functions of higher types has a steeper curve. Note that this subsumes the assumption we made about the cost curves in the previous example (there high types has a flatter curve, but $c(x, \theta) = -v(x, \theta)$, so this goes through.)

For the rest of this document, we'll assume that the Regularity Conditions of Milgrom and Segal's Envelope Theorem are satisfied by v. These are pretty standard assumptions, and they hold in practically every example.

Now, consider a direct mechanism $\delta = (x(.), t(.))$. Define $V(\theta) := v(x(\theta), \theta) + t(\theta)$ as the payoff from truth-telling, given mechanism δ. We're looking for conditions to guarantee that δ is incentive compatible:

$$\theta \in \arg\max_{\theta' \in \Theta} v(x(\theta'), \theta) + t(\theta') \quad \forall \theta \in \Theta \qquad (IC)$$

We'll present two conditions. The first one is the local incentive compatibility:

$$V(\theta) = V(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} v_\theta(x(s), s) ds \quad \forall \theta \in \Theta \qquad (ICFOC)$$

Realize that this condition has the feature that it pins down the transfers $t(.)$ given the allocation rule $x(.)$. The second condition is monotonicity:

$$x(\theta) \text{ is nondecreasing in } \theta. \qquad (MON)$$

Here comes the critical theorem:

Theorem 1. ***(Constraint Simplification.)*** *Suppose v has SID. Then*

$$(IC) \Leftrightarrow (ICFOC) + (MON)$$

Take a moment to appreciate how much easier this theorem will make our lives! I'll not present the proof here because it would be a one-by-one rip-off from Juuso's lecture notes, but you should surely check it. (Designing Optimal Mechanisms, p.10.)

Also realize that, since (ICFOC) helps us pin down the transfer rule, this theorem tells us; "don't worry about incentive compatibility as long as $x(.)$ is monotone, as you can always find $t(.)$ to make this allocation incentive compatible." In other words, we get the following corollary, which I presented in the previous section, for free:

[11] There obviously is a relation between SID and Strict Single Crossing Property (SSCP). To see the formulation which gives exact equivalence, check Juuso's lecture notes (Designing Optimal Mechanisms, p.8.)

Corollary 1. *Suppose v has SID. Then $x(.)$ is implementable if and only if it is nondecreasing.*

3.3 The General Case

As argued above, once the necessary armament is built, the solution of a standard screening problem is pretty straightforward. In particular, once the problem is specified (with (IC) and (IR) constraints for all types) and SID is assumed, these are the steps that one needs to follow:

1. Use the Constraint Simplification Theorem to replace (IC) with (ICFOC) and (MON).
2. Argue that, only the lowest type's (IR) constraint should matter.
3. Substitute (ICFOC) into the objective function and change the order of integration to derive the virtual values.
4. Ignore (MON) and solve pointwise, hope that (MON) holds at the optimal solution to the relaxed problem.
5. If (MON) holds at the optimal solution, done! If it fails, iron!

As an illustrative example, let's consider the general version of the procurement problem, where $\Theta = [\underline{\theta}, \bar{\theta}]$ instead of only two types. Suppose that θ is distributed with density f (cdf F). We will retain the assumptions about U and C. This implies that in this context we will have $V = -c(x, \theta) + t$; in other words, $-c(x, \theta)$ will work as $v(x, \theta)$ in this example. Note that, then, the assumption of $C_{x\theta} < 0$ guarantees that $v_{x\theta} > 0$, i.e. SID is satisfied. Also for future reference, realize that $v_\theta = -C_\theta$. Finally, $V(\theta) = -C(x(\theta), \theta) + t(\theta)$.

For simplicity, we will assume that $U(x) = x$. Then the problem becomes:

$$\begin{aligned}
&\max_{(x(.),t(.))} && \int_{\underline{\theta}}^{\bar{\theta}} (x(\theta) - t(\theta)) f(\theta) d\theta && \\
&\text{s.t.} && \theta \in \arg\max_{\theta'} \left[-C(x(\theta'), \theta) + t(\theta')\right] && \forall \theta \in \Theta \quad (IC) \\
& && -C(x(\theta), \theta) + t(\theta) \geq 0 && \forall \theta \in \Theta \quad (IR)
\end{aligned}$$

Let's follow the steps above. By SID, one can replace (IC) with (ICFOC) and (MON). Furthermore, it's easy to see that (IC), along with the (IR) for the lowest type, implies (IR) for all types, so it suffices to keep (IR) for lowest type only. Thus the problem becomes:

$$\begin{aligned}
&\max_{(x(.),t(.))} && \int_{\underline{\theta}}^{\bar{\theta}} (x(\theta) - t(\theta)) f(\theta) d\theta && \\
&\text{s.t.} && -C(x(\theta), \theta) + t(\theta) = -C(x(\underline{\theta}), \underline{\theta}) + t(\underline{\theta}) - \int_{\underline{\theta}}^{\theta} C_\theta(x(s), s) ds \quad \forall \theta \in \Theta && (ICFOC) \\
& && x(\theta) \text{ is nondecreasing in } \theta && (MON) \\
& && C(x(\underline{\theta}), \underline{\theta}) + t(\underline{\theta}) \geq 0 && (IR)
\end{aligned}$$

It's easy to see that (IR) for the lowest type must bind at the optimum (otherwise decrease $t(\underline{\theta})$). Plug this back into (ICFOC) to get: $t(\theta) = C(x(\theta), \theta) - \int_{\underline{\theta}}^{\theta} C_\theta(x(s), s) ds$ for each θ. Finally, substitute this into the objective function to simplify the program into:

$$\begin{aligned}
&\max_{x(.),t(\underline{\theta})} && \int_{\underline{\theta}}^{\bar{\theta}} \left[x(\theta) - C(x(\theta), \theta) + \int_{\underline{\theta}}^{\theta} C_\theta(x(s), s) ds \right] f(\theta) d\theta && \\
&\text{s.t.} && x(\theta) \text{ is nondecreasing in } \theta && (MON) \\
& && t(\underline{\theta}) = C(x(\underline{\theta}), \underline{\theta}) && (IR)
\end{aligned}$$

Now, realize that the objective function can be written as:

$$\begin{aligned}
\int_{\underline{\theta}}^{\bar{\theta}} \left(x(\theta) - C(x(\theta),\theta) + \int_{\underline{\theta}}^{\theta} C_\theta(x(s),s)ds \right) f(\theta)d\theta &= \int_{\underline{\theta}}^{\bar{\theta}} [x(\theta) - C(x(\theta),\theta)]\, f(\theta)d\theta + \int_{\underline{\theta}}^{\bar{\theta}}\int_{\underline{\theta}}^{\theta} C_\theta(x(s),s)f(\theta)dsd\theta \\
&= \int_{\underline{\theta}}^{\bar{\theta}} [x(\theta) - C(x(\theta),\theta)]\, f(\theta)d\theta + \int_{\underline{\theta}}^{\bar{\theta}}\int_{s}^{\bar{\theta}} C_\theta(x(s),s)f(\theta)d\theta ds \\
&= \int_{\underline{\theta}}^{\bar{\theta}} [x(\theta) - C(x(\theta),\theta)]\, f(\theta)d\theta + \int_{\underline{\theta}}^{\bar{\theta}} C_\theta(x(s),s) \left(\int_{s}^{\bar{\theta}} f(\theta)d\theta \right) ds \\
&= \int_{\underline{\theta}}^{\bar{\theta}} [x(\theta) - C(x(\theta),\theta)]\, f(\theta)d\theta + \int_{\underline{\theta}}^{\bar{\theta}} C_\theta(x(s),s)(1-F(s))ds \\
&= \int_{\underline{\theta}}^{\bar{\theta}} [x(\theta) - C(x(\theta),\theta)]\, f(\theta)d\theta + \int_{\underline{\theta}}^{\bar{\theta}} C_\theta(x(\theta),\theta)(1-F(\theta))d\theta \\
&= \int_{\underline{\theta}}^{\bar{\theta}} \left[x(\theta) - C(x(\theta),\theta) + C_\theta(x(\theta),\theta)\frac{1-F(\theta)}{f(\theta)} \right] f(\theta)d\theta \\
&= \mathbb{E}\left[x(\theta) - C(x(\theta),\theta) + C_\theta(x(\theta),\theta)\frac{1-F(\theta)}{f(\theta)} \right]
\end{aligned}$$

We can now see that the infamous "virtual cost" $\hat{C}(x,\theta) := C(x,\theta) - C_\theta(x,\theta)\frac{1-F(\theta)}{f(\theta)} \geq C(x,\theta)$ appears in the objective function. The interpretation is the same as the one provided previously. Importantly, the problem now becomes:

$$\begin{aligned}
\max_{x(.),t(\underline{\theta})} \quad & \mathbb{E}\left[x(\theta) - \hat{C}(x(\theta),\theta) \right] & \\
\text{s.t.} \quad & x(\theta) \textit{ is nondecreasing in } \theta & (MON) \\
& t(\underline{\theta}) = C(x(\underline{\theta}),\underline{\theta}) & (IR)
\end{aligned}$$

Which, if one ignores (MON), can be solved pointwise! In particular, the pointwise solution to the relaxed problem gives the first-order condition which defines $x^*(\theta)$ for each θ:

$$\hat{C}_x(x^*(\theta),\theta) = 1 \Leftrightarrow C_x(x^*(\theta),\theta) - C_{x\theta}(x^*(\theta),\theta)\frac{1-F(\theta)}{f(\theta)} = 1$$

One can right away see from this equation that many properties of the two-type case are retained. In particular, the highest type's allocation is efficient ($F(\bar{\theta}) = 1$), and there are downward distortions for other types.

One can now check whether $x^*(\theta)$ satisfies (MON). If the distribution of θ, f, is "regular" enough, it should be monotone (in many cases it will be). If not, one needs to do some ironing. When it comes to ironing, though, it helps to remember Bengt's wisdom:

> "Myerson said that it took one week for him to figure out what to do when $x^*(.)$ is not monotone. And when it comes to Myerson vs common folks, the dog years-human years conversion applies. If it took one week for Myerson to iron a function, it will take seven years for you! Long story short: Don't get into ironing."

If you're really curius about the basics of ironing though, Jason Hartline's book (Mechanism Design and Approximation, Chapter 3) is a useful resource to understand the basics.

So, that's all for screening! Virtually all applications in this topic follow the same methodology. Mussa and Rosen (1978) does this in the context of a monopolist determining product quality, Maskin and Riley (1984) does this in the context of nonlinear pricing, Baron and Myerson (1982) does this in the context of regulation, Laffont and Tirole (1986) does this in the context of procurement... In one way or another, these papers essentially do the same thing once the problem is stated clearly.

Ah, and then there is Myerson (1981), which does this in the context of auctions! Who can forget it? Best paper ever.

3.4 Optimal Auctions

We are in the auctions setup, but this is pretty much a specific case of the general setup given above, with $v_i(x, \theta_i) = \theta_i x_i$ (which, unsurprisingly, satisfies SID). Juuso has a nice description and discussion of the model in the lecture notes (Designing Optimal Mechanisms), even though the paper itself is pretty readable too. It's a fun and useful exercise to check that solution follows the steps above, exactly. In the end, the "virtual value" of agent with type θ_i turns out to be: $\theta_i - \frac{1-F_i(\theta_i)}{f_i(\theta_i)}$. All in all, the simplified problem becomes:

$$\begin{aligned} \max_{x(.)} \quad & \mathbb{E} \ \ \theta_0 x_0(\theta) + \sum_{i=1}^{N} x_i(\theta)(\theta_i - \frac{1-F_i(\theta_i)}{f_i(\theta_i)}) \\ \text{s.t.} \quad & \bar{x}_i(\theta_i) \ \textit{is nondecreasing in} \ \theta_i \qquad\qquad (MON) \end{aligned}$$

Which can be solved pointwise to show that the optimal auction gives the good to the bidder with highest virtual value, i.e. it runs an auction on virtual values (rather than values). Of course, if the highest virtual value belongs to the seller, she would keep the good to herself. This means that if $\theta_i - \frac{1-F_i(\theta_i)}{f_i(\theta_i)} < \theta_0$ for each i, the good will not be sold. Realize that there might still be cases where $\theta_i > \theta_i$, i.e. it is indeed efficient to sell the good, yet the seller still doesn't sell it. In other words, the allocation is not guaranteed to be efficient.[12] But this is hardly surprising: we said that the auction is revenue-maximizing, not efficient. If you want efficiency, run VCG, i.e. run a second-price auction! If you want optimality, give the good to the buyer with highest virtual value. Under regular distributions and symmetric bidders, this becomes equivalent to running a second-price auction with reserve price. By putting a reserve price, the seller risks not selling the good when it is sometimes profitable to do so (i.e. when everyone bids below the reserve price), but as a compensation receives a higher transfer in some cases (when only one bidder bids above the reserve price.) The reserve price is set such that there is a balance in favor of the latter scenario.

A more important thing to realize is that, in the simplified problem above, the transfers son't appear at all. This gives way to the famous Revenue Equivalence Theorem, which is indeed nothing more than a restatement of the Payoff Equivalence Theorem for BIC mechanisms.

Theorem 2. ***(Revenue Equivalence Theorem.)*** *Consider any two BIC mechanisms $\delta = (x(.), t(.))$ and $\hat{\delta} = (x(.), \hat{t}(.))$. If $V_i(\underline{\theta}_i) = \hat{V}_i(\underline{\theta}_i)$ for all $i = 1, \ldots, N$, then the seller's expected profit is the same under both mechanisms.*
Corollary 2. *"Standard Auctions" (FPA, SPA, English, Dutch, All-Pay...) yield the same expected revenues to the auctioneer.*

To see why the corollary holds, it suffices to see that all formats have equilibria in strictly increasing bidding strategies (thus the winner is the same, and the allocation rule is the same).

This is a truly powerful result, and it's quite popular too! Indeed I vividly remember Parag saying that "find bidding strategies in [place auction format here]" was a standard theory general question in Harvard when he was a student (which is not distant past), and the trick is clearly using revenue equivalence. I'm adding a question along those lines that Juuso asked in 14.281 PSet 1 Fall 2014, and my answer to that question, so that it gives an idea. (The website involves the original solution – I'm adding mine not because I find it bette than the original one but because I have the Tex code for mine.) I'm also attaching a question asked in last year's theory general by Parag, which (I think) should be solved in the same manner.

[12]Depending on the distributions, the auction may give the object to an agent who doesn't have the highest value, either.

4 Appendix

Question 2. Auctions with iid buyers, and where the seller has zero valuation for the good.

Suppose the buyers have privately known valuation over [0,1], where $\theta_i \in [0,1]$ (the valuation of buyer i) is i.i.d. with according to some F on $[0,1]$, and F is common across all buyers.
For notational simplicity, let me define the random variable x, which is distributed according to $G := F^{n-1}$, which denotes the maximum valuation among the remaining bidders. Note that the pdf of x is given by $g(x) = (n-1)(F(x))^{n-2}f(x)$.

(a) In a standard second-price auction, the interim payment of a buyer of type θ_i is given by:

$$\begin{aligned}
\bar{\psi}_i(\theta_i) &= Pr\{\theta_i \geq \theta_j \forall j \in \{1,\dots,n\}\}\mathbb{E}[\theta_{(2)}|\theta_{(1)} = \theta_i] \\
&= Pr\{x \leq \theta_i\}\mathbb{E}[x|x \leq \theta_i] \\
&= G(\theta_i)\mathbb{E}[x|x \leq \theta_i] \\
&= (F(\theta_i))^{n-1} \int_0^{\theta_i} x\frac{(n-1)(F(x))^{n-2}f(x)}{(F(\theta_i))^{n-1}}dx \\
&= \int_0^{\theta_i} x(n-1)(F(x))^{n-2}f(x)dx
\end{aligned}$$

And the government's expected revenue is simply the expected value of the second highest valuation ($\mathbb{E}[\theta_{(2)}]$). As we know from the distributions of order statistics, the pdf of $\theta_{(2)}$ is given by the function: $f_{(2)}$, where:

$$f(x) = n(n-1)(F(x))^{n-2}(1-F(x))f(x)$$

So we have:

$$\mathbb{E}[\theta_{(2)}] = \int_0^1 n(n-1)x(F(x))^{n-2}(1-F(x))f(x)dx$$

Realize that, by payoff equivalence, these are the expected payments and revenues in any standard auction (first-price, second-price, all-pay etc.) so one can easily use these equations to calculate the optimal bidding strategies in a BNE under these auctions (as we'll do in the second part.)

(b) The "lobbying game" involves the potential buyers engaging in efforts *before the decision is made*, which are clearly non-refundable and irreversible. If we assume that the committee awards the highest effort, then this effectively turns the problem into an all-pay auction (where the good is awarded to the highest bidder, and the bids are non-refundable and irreversible.)[13]
So what are the BNE strategies in an all-pay auction? We know that both the second-price auction and the all-pay auction are "standard" auctions: they allocate the good to the highest-valuation buyer, so they have the same allocation rule. By payoff equivalence, then, interim payments must be equal (indeed they can differ by a constant but it's easy to see that type-0 buyer bids 0 in both auctions, so the constant has to be equal to zero.) We've just derived that in a second-price auction, interim payment of an agent with valuation θ_i is: $\int_0^{\theta_i} x(n-1)(F(x))^{n-2}f(x)dx$. In an all-pay auction, if the agent bids $b(\theta_i)$, and pays it surely, so the interim payment is: $b(\theta_i)$. Setting them equal, we get:

$$b(\theta_i) = \int_0^{\theta_i} x(n-1)(F(x))^{n-2}f(x)dx$$

Notice that we can get a much cleaner result if we assume that F is the uniform distribution over [0,1]. In particular, we know that $F(\theta_i) = \theta_i$ if $\theta_i \sim U[0,1]$, and that the expected values of the order

[13] The idea of modeling the lobbying game as an all-pay auction dates back to Grossman and Helpman (1994). All-pay auctions indeed arguably correspond to many other realistic cases, for instance, sports contests, as argued in Moldovanu and Sela (2001).

statistics divide the line evenly in a uniform distribution, such that $\mathbb{E}[\theta_{(2)}|\theta_{(1)} = \theta_i] = \frac{n-1}{n}$. This leads us to the result:

$$\begin{aligned} b(\theta) = (F(\theta_i))^{n-1} &\int_0^{\theta_i} x\frac{(n-1)(F(x))^{n-2}f(x)}{(F(\theta_i))^{n-1}}dx \\ &= (F(\theta_i))^{n-1}\mathbb{E}[\theta_{(2)}|\theta_{(1)} = \theta_i] \\ &= (\theta)^{n-1}[\theta\frac{n-1}{n}] \\ &= (\theta)^n\frac{n-1}{n} \end{aligned}$$

in an all-pay auction with uniform distribution of values over $[0,1]$.
Realize that as n grows, $b(\theta)$ indeed pointwise converges to the step function. This means that the bid shading motive dominates for a quite large range of θ (for $\theta \leq 1-\varepsilon$), simply because with more bidders a buyer has a smaller probability of winning, which reduces the incentives to bid more aggressively. Only for large θ ($> 1-\varepsilon$, where $\varepsilon \to 0$), the equilibrium bid goes to one, since this type almost surely wins the auction.

(c) As we derived in class, in an optimal auction, the seller gives the object to the agent with highest virtual value, defined as $\phi(\theta) = \theta - \frac{1-F(\theta)}{f(\theta)}$. Remember that the uniform distribution is within the class of "regular" distributions, in the sense that $\phi(\theta)$ is increasing in θ, so that the optimal auction can be implemented in a simple manner, say, for instance, as a second-price auction with reserve price. With F uniform over $[0,1]$, the optimal reserve price (θ^{res} such that $\theta^{res} = \frac{1-F(\theta^{res})}{f(\theta^{res})}$) turns out to be $\theta^{res} = \frac{1}{2}$. The revenue in an optimal auction is:

$$Pr\{\theta_{(1)} < \frac{1}{2}\}0 + Pr\{\theta_{(2)} < \frac{1}{2} < \theta_{(1)}\}\frac{1}{2} + Pr\{\frac{1}{2} < \theta_{(2)}\}\mathbb{E}[\theta_{(2)}|\frac{1}{2} < \theta_{(2)}]$$

Whereas the expected revenue in a second price auction without reserve price is simply $\mathbb{E}[\theta_{(2)}]$. With uniform distribution and with $n+1$ buyers, this equals: $\frac{n-1}{n+1}$.
Our claim is that the latter is larger than the former, i.e. efficient auction with $n+1$ buyers yields a higher revenue than the optimal auction with n buyers. One can explicitly calculate these quantities (which is an arduous task at best) and see that it holds, but it's worth emphasizing that the results holds with general distributions as well. This is the well-known Bulow-Klemperer result, and here is a simple proof (based on Hartline's lectures, and is first presented by Kirkegaard (2006) as far as I see.)
Proposition 1. *The efficient auction with $n+1$ buyers yields a higher expected revenue that the optimal auction with n buyers.*

Proof. As a shorthand notation, I'll refer to the expected revenue of efficient auction with k buyers as EFF_k, and that of the optimal auction as OPT_k. We need to show that $EFF_{n+1} \geq OPT_n$. Now define the alternative mechanism for k buyers, which in shorthand I'll refer to as ALT, as follows:

- Take an arbitrary subset of $k-1$ buyers and run the optimal mechanism (second price auction with reserve price) among these buyers.
- If the item is sold at Step 1, done. If not, give the item to the k'th agent (who is left outside at Step 1) for free.

The first thing is to realize that $ALT_{n+1} \geq OPT_n$, and this is by construction (since ALT_{n+1} runs OPT_n at Step 1, it can't do worse than that.)[14] The second thing is to realize that the auction format ALT indeed resides within the set of *auction formats in which the good is sold with probability one*. Define this set of auction formats as $\mathcal{M}$. Realize that EFF is also an element of $\mathcal{M}$, since it sells the good with probability one. Moreover, our claim is that EFF is the revenue-maximizing rule

[14]Of course, the iid'ness of distributions and IPV assumption is crucial here.

within this set. For the sake of this claim, though, it suffices to show that EFF_{n+1} yields a higher revenue than ALT_{n+1}. But this is easy to see, since there is always a range of valuations in which EFF yields a higher revenue than ALT, whereas the opposite case is not possible. We conclude that $EFF_{n+1} \geq ALT_{n+1}$. Combining both observations yield $EFF_{n+1} \geq OPT_n$.[15] □

So, what does this result tell us? Intuitively it says that what matters in an auction is competition, rather than clever design. Therefore, if it were the government which runs the auction, it should try to attract as many potential buyers as possible, rather than focusing on figuring out what the optimal auction format is. (Of course, there is no harm in trying both! The theorem just tells that the first goal is more fruitful than the second.)

(d) As we derived in the lecture, the reserve price should be such that any type $\theta < \theta^{res}$, where θ^{res} is defined such that $\theta^{res} = \frac{1-F(\theta^{res})}{f(\theta^{res})}$, does not want to bid at all. In a first price auction, the only way to guarantee this is to set the reserve price to θ^{res} (it's easy to see that for a lower cutoff some other types $\theta < \theta^{res}$ will bid a strictly positive amount and will have a strictly positive probability to win; for a higher cutoff, some types $\theta > \theta^{res}$ are drawn out of the auction, yielding a lower revenue.)
For the uniform case, therefore, the government needs to set $b^{min} = \theta^{res} = \frac{1}{2}$. Clearly, for any $\theta_i < \frac{1}{2}$, the optimal bid is equal to zero. For types higher than $\frac{1}{2}$, we can use payoff equivalence to derive the equilibrium bids. Remember that we are in the regular case, thus the mechanism always allocates the good to the highest-value agent.[16] Therefore we have, for any type $\theta_i > 1/2$, $\bar{x}(\theta_i) = Pr\{\theta_i = \theta_{(1)}\} = (\theta_i)^{n-1}$.
Note that type θ_i's interim expected payoff, when she bids $b(\theta_i)$, is equal to $(\theta_i - b(\theta_i))\bar{x}(\theta_i)$. By the payoff equivalence theorem, we know that the interim payoff is equal to:

$$V(\theta_i) = V(\frac{1}{2}) + \int_{1/2}^{\theta_i} \bar{x}(s)ds = \int_{1/2}^{\theta_i} (s)^{n-1}ds$$

Equating the two expressions, we obtain:

$$(\theta_i - b(\theta_i))\bar{x}(\theta_i) = \int_{1/2}^{\theta_i} (s)^{n-1}ds$$

Which, by simple algebra, yields: $b(\theta) = \theta - \frac{\theta^n - (1/2)^n}{n\theta^{n-1}}$. Thus the equilibrium bidding function is:

$$b(\theta) = \begin{cases} 0, & \text{if } \theta \leq 1/2; \\ \theta - \frac{\theta^n - (1/2)^n}{n\theta^{n-1}}, & \text{if } \theta \geq 1/2. \end{cases}$$

For the all-pay auction, we can use the same trick to calculate the minimum acceptable contribution. Realize that we still have $\bar{x}(\theta_i) = (\theta_i)^{n-1}$ for $\theta_i \geq \frac{1}{2}$. The key observation here is that type 1/2 must be indifferent between bidding (which yields an interim utility of $\frac{1}{2}\bar{x}(\theta_i) - b(1/2)$) and not bidding (which yields a payoff of zero.) Equating, we obtain: $b(1/2) = \frac{1}{2}\bar{x}(\frac{1}{2}) = \frac{1}{2}(\frac{1}{2})^{n-1} = (\frac{1}{2})^n$. Since we want any buyer with value $\theta_i \geq \frac{1}{2}$ to enter, the minimum acceptable contribution should be: $b^{min} = b(1/2) = (\frac{1}{2})^n$.

(e) Now consider the case with divisible assets, with agent of type θ_i having utility: $\theta_i v(x_i) - t_i$.

(i) The first-best allocation rule, where types $\theta = (\theta_1, \ldots, \theta_n)$ are observable, is the solution to:

$$\max_{x=(x_1,\ldots,x_n)} \sum_i \theta_i v_i(x_i)$$

subject to

$$\sum_i x_i \leq 1$$

[15] It's also easy to see that the inequality is indeed strict.
[16] That is, conditional on agent's valuation being above 1/2.

It's easy to see that the constraint must bind at the optimum, since v_i is strictly increasing. Then, denoting λ as the Lagrange multiplier associated with the constraint, one obtains the first order condition:

$$\theta_i v'(x_i) = \lambda \text{ for each } i.$$

In other words, the allocation is such that the marginal utilities are equalized across agents. With $v_i(x_i) = 2\sqrt{x_i}$, this corresponds to: $\frac{\theta_i}{\sqrt{x_i}} = \lambda$, or equivalently: $x_i = \frac{(\theta_i)^2}{\lambda^2}$ for each i. Because the feasibility constraint must bind, we have:

$$x_i = \frac{(\theta_i)^2}{\sum_j (\theta_j)^2} \text{ for each } i.$$

i.e. each agent's share is proportional to her type-squared.

(ii) When types are private information, the optimal BIC mechanism $\mu = (x, t)$ is the solution to:

$$\max_{x(.),t(.)} \mathbb{E}[\sum_i t_i(\theta_i)]$$

subject to

$$(IC_{\theta_i}) \forall i, \theta_i$$
$$\theta_i v_i(x_i(\theta_i)) - t_i(\theta_i) \geq 0 \forall i, \theta_i$$
$$\sum_i x_i(\theta_i) \leq 1 \forall \theta$$

Standard arguments apply: one could replace (IC) with $(ICFOC) + (MON)$, and it's easy to show that the individual rationality constraint binds for the lowest type only. By $(ICFOC)$, we obtain:

$$\bar{t_i}(\theta_i) = \theta_i v_i(\bar{x}_i(\theta_i)) - \int_0^{\theta_i} v_i(\bar{x}_i(s))ds + t_i(0)$$

One can substitute this into the objective function to simplify the problem into:

$$\max_{x(.),t(.)} \mathbb{E}[\sum_i \theta_i v_i(\bar{x}_i(\theta_i)) - \int_0^{\theta_i} v_i(\bar{x}_i(s))ds + t_i(0)]$$

subject to

$$(MON)$$
$$t_i(0) \leq 0 \forall i$$
$$\sum_i x_i(\theta_i) \leq 1 \forall \theta$$

Clearly, $t_i(0) = 0$ at the optimal solution. Ignoring (MON) and changing the order of integration translates the problem into:

$$\max_{x(.)} \mathbb{E}[\sum_i \theta_i v_i(x_i(\theta_i))(\theta_i - \frac{1 - F(\theta_i)}{f(\theta_i)}]$$

subject to

$$\sum_i x_i(\theta_i) \leq 1 \forall \theta$$

Again it's easy to show that the constraint must bind at the optimum, thus denoting its Lagrange multiplier as λ yields the interior first-order condition:

$$v_i'(x_i(\theta_i))(\theta_i - \frac{1-F(\theta_i)}{f(\theta_i)}) = \lambda \text{ for each } i.$$

(Clearly, if $\theta_i - \frac{1-F(\theta_i)}{f(\theta_i)} < 0$, we have the boundary solution $x_i(\theta_i) = 0$.) In other words, the allocation is such that the marginal utilities times the virtual valuations are equalized across agents. With $v_i(x_i) = 2\sqrt{x_i}$, this corresponds to: $\frac{1}{\sqrt{x_i}}(\theta_i - \frac{1-F(\theta_i)}{f(\theta_i)}) = \lambda$, or equivalently: $x_i = \frac{[\theta_i - \frac{1-F(\theta_i)}{f(\theta_i)}]^2}{\lambda^2}$ for each i. Because the feasibility constraint must bind, we have:

$$x_i = \frac{\phi_i(\theta_i)^2}{\sum_j (\phi_j(\theta_j))^2} \text{ for each } i$$

where $\phi_i(\theta_i) = \max\{\theta_i - \frac{1-F_i(\theta_i)}{f_i(\theta_i)}, 0\}$.
i.e. each agent's share is proportional to her virtual value-squared. Again, under the regular case, the allocation rule is monotonic, so this is indeed the optimal rule.

(iii) The undesirability of randomized mechanisms is a natural consequence of the concavity of $v_i(.)$ in this setup. Since $v_i(.)$ is concave, for any randomized mechanism, the government would simply be better off by offering the expectation of the allocation rule, which will still be feasible, but will generate a higher revenue because now the agent is willing to pay more. This is because the agent's willingness to pay is: $v_i(\mathbb{E}[X_i]) \geq \mathbb{E}[v_i(X_i)]$ by concavity of v and Jensen's inequality, thus the government can extract more surplus when it replaces the randomized allocation rule with the expectation-equivalent one.

Question 3. Bilateral trade as in Myerson-Satterthwaite.

(a) Take $\Theta_1 = \{0, 2\}$ and $\Theta_2 = \{1\}$. Then the efficient allocation rule x satisfies: $x(0, 1) = 1$ and $x(2, 1) = 0$ (the good is sold and not sold, respectively). We can simply set $t(0, 1) = t \in (0, 1)$ and $t(2, 1) = 0$ (this is the amount the buyer has to pay to the seller). Then both players receive non-negative utility, trading rule is efficient, incentive compatibility is satisfied, and budget is balanced. Obviously the same reasoning works also when the buyer's type space has more than one element, say $\Theta_2 = \{1, 3\}$, and when each of the buyer's and seller's types is equally likely (we are assuming independence, of course), in which case the efficient rule can be implemented by the transfer scheme: $t(0, 1) = 1, t(0, 3) = 2, t(2, 1) = 0, t(2, 3) = 3$.

(b) Now, the utilities of the players are $U_1(x, t_1, \theta_1) = (1-x-y)\theta_1 + t_1$ and $U_2(x, t_2, \theta_2) = t_2 - (1-x-y)\theta_2$, where x is the efficient trading rule which indicates whether or not $\theta_2 \geq \theta_1$. This allows us to calculate the social surplus (sum of utilities) from the efficient trading rule as $S(\theta) = (\theta_2 - \theta_1)^+ - (\theta_2 - \theta_1)(1-y)$.

Now, we know that, under BIC, Groves scheme is efficient, IC and IR, but it runs a deficit of $(n-1)\mathbb{E}[S(\theta)]$, where S is the expected social surplus under the efficient mechanism. If we let $\mathbb{E}[S(\theta)|\theta_i]$ denote the interim expected social surplus given i observing her type θ_i, we know that we can subtract this from the other players' transfers since it only depends on θ_i.[17] However, we need to make sure that even the "weakest" players obtain non-negative expected utility, which is why we can subtract at most $\sum_i \min_{\theta_i} \mathbb{E}[S(\theta)|\theta_i]$. Therefore, we can achieve BB if and only if the following condition holds:

$$\sum_i \min_{\theta_i} \mathbb{E}[S(\theta)|\theta_i] \geq (n-1)\mathbb{E}[S(\theta)]$$

Because we are assuming that the types follow independent standard uniform distributions, we can calculate:

[17] Note that this is what AGV does.

1. $E[S(\theta)] = E[(\theta_2 - \theta_1)^+] = \frac{1}{2}\frac{1}{3} = \frac{1}{6}$,
2. $\mathbb{E}[S(\theta)|\theta_1] = (1-\theta_1)\frac{1-\theta_1}{2} + (1-y)\theta_1 + \frac{y}{2} \Rightarrow \min_{\theta_1} \mathbb{E}[S(\theta)|\theta_1] = \frac{y}{2}(1-y)$,
3. $\mathbb{E}[S(\theta)|\theta_2] = \frac{\theta_2^2}{2} - (1-y)\theta_2 + \frac{1-y}{2} \Rightarrow \min_{\theta_2} \mathbb{E}[S(\theta)|\theta_2] = \frac{y}{2}(1-y)$

$\Rightarrow y(1-y) \geq \frac{1}{6}$ which is equivalent to $y \in \left[\frac{3-\sqrt(3)}{6}, \frac{3+\sqrt(3)}{6}\right]$.