Models in Microeconomic Theory

Martin J. Osborne
Ariel Rubinstein

https://www.openbookpublishers.com

ISBN Paperback: 978-1-78374-920-1
ISBN Hardback: 978-1-78374-921-8
ISBN Digital (PDF): 978-1-78374-922-5
DOI: 10.11647/OBP.0204

This book is available in two versions, one that uses feminine pronouns and one that uses masculine pronouns. This version uses masculine pronouns.

Version: 2020.12.21 (h)

17 Mechanism design

The models in Parts II and III analyze the behavior of individuals given a specific structure for their interaction. In this chapter, we turn this methodology on its head. That is, we seek a set of rules for the interaction between the individuals that generates specific outcomes. Analyses of this type are called "mechanism design". This field is huge; we demonstrate some of the basic ideas through a simple model.

17.1 Deciding on a public project

A community of individuals has to decide whether to carry out a joint project. For example, the inhabitants of a city consider building a new subway, or the tenants in a neighborhood consider adding a bench to their community garden. The action to be taken is public in the sense that all individuals are affected by it. The community can either undertake the project or not.

Definition 17.1: Public project problem

A *public project problem* $\langle N, D\rangle$ consists of a set $N = \{1, \ldots, n\}$ of individuals and a set $D = \{0, 1\}$ of public decisions (1 means a project is executed and 0 means it is not).

The individuals may differ in their attitudes to the project: some may support it and some may oppose it. We look for mechanisms that balance these interests. The mechanism is allowed to require the agents to make and receive payments, which are used to induce a desirable outcome. The presence of payments mean that the individuals' preferences have to be defined not on the set D but on pairs of the type (d, t) where d is the public decision and t is the transfer (positive, zero, or negative) to the individual.

Each individual i is characterized by a number v^i, which may be positive or negative, with the interpretation that he is indifferent between (*i*) the project's not being carried out and his not making or receiving any payment and (*ii*) the project's being carried out and his paying v^i (when $v^i > 0$) or receiving $-v^i$ (when $v^i < 0$). Thus if $v^i > 0$ then i benefits from the project and is willing to pay up to v^i to have it realized; if $v^i < 0$ then i is hurt by the project and is willing to

Chapter of *Models in Microeconomic Theory* by Martin J. Osborne and Ariel Rubinstein. Version 2020.12.21 (h).
 https://doi.org/10.11647/OBP.0204.17

pay up to $-v^i$ to stop its being realized. If $v^i = 0$ then i is indifferent between the project's being executed and not. Note that the interests of an individual depend only on his own valuation of the project.

Definition 17.2: Valuation profile

A *valuation profile* $(v^i)_{i \in N}$ for a public project problem $\langle N, D \rangle$ consists of a number v^i for each individual $i \in N$. The number v^i determines i's preferences over pairs (d, t^i) consisting of a public decision $d \in D$ and a number t^i, the amount of money transferred to (if $t^i > 0$) or from (if $t^i < 0$) individual i. Specifically, the preferences of each individual $i \in N$ over pairs (d, t^i) with $d \in D$ and $t^i \in \mathbb{R}$ are represented by the utility function

$$\begin{cases} v^i + t^i & \text{if } d = 1 \\ t^i & \text{if } d = 0. \end{cases}$$

17.2 Strategy-proof mechanisms

We assume that the valuation v^i of each individual i is known only to him. So if the community wants to base its decision on these valuations, it needs to query the individuals. A direct mechanism with transfers asks each individual to report a number and interprets this number as his valuation. The mechanism then specifies the public decision and the monetary transfers to or from the individuals, as a function of their reports.

Definition 17.3: Direct mechanism with transfers

For a public project problem $\langle N, D \rangle$, a *direct mechanism with transfers* is a collection $(\delta, \tau^1, \ldots, \tau^n)$ of functions that assign to each profile $(x^1, \ldots, x^n)$ of numbers (the individuals' reports) a public decision $\delta(x^1, \ldots, x^n) \in D$ and a monetary transfer $\tau^i(x^1, \ldots, x^n)$ for each $i \in N$.

Each individual can report any number he wishes, so we need to consider the possibility that individuals may benefit from reporting numbers different from their valuations. Intuitively, an individual may benefit from exaggerating his valuation of the project positively if he supports it and negatively if he opposes it. We say that a mechanism is strategy-proof if every individual, whatever his valuation, optimally reports this valuation, regardless of the other individuals' reports. That is, given any reports of the other individuals, no individual can do better than reporting his valuation.

Definition 17.4: Strategy-proof mechanism

For a public project problem $\langle N, D\rangle$, a direct mechanism with transfers $(\delta, \tau^1, \ldots, \tau^n)$ is *strategy-proof* if for every valuation profile $(v^i)_{i\in N}$, every individual $i \in N$, every list $(x^1, \ldots, x^{i-1}, x^{i+1}, \ldots, x^n)$ of numbers (reports of the other individuals), and every number z^i (report of i) we have

$$\delta(x^1, \ldots, v^i, \ldots, x^n)v^i + \tau^i(x^1, \ldots, v^i, \ldots, x^n) \geq \delta(x^1, \ldots, z^i, \ldots, x^n)v^i + \tau^i(x^1, \ldots, z^i, \ldots, x^n).$$

That is, i optimally reports his valuation, whatever it is, regardless of the other individuals' reports.

Notice that the definition does not require that an individual's true valuation is the only optimal report for him regardless of the other individuals' reports.

An example of a strategy-proof mechanism is majority rule.

Example 17.1: Majority rule

Majority rule is the direct mechanism in which the project is executed if and only if a majority of individuals report a positive number, and no monetary transfers are made. That is,

$$\delta(x^1, \ldots, x^n) = 1 \text{ if and only if } |\{i \in N : x^i > 0\}| > n/2$$

and $\tau^i(x^1, \ldots, x^n) = 0$ for all $i \in N$ and for all profiles $(x^1, \ldots, x^n)$.

This mechanism is strategy-proof. Take an individual with a positive valuation. His changing his report from one positive number to another has no effect on the outcome. His switching from a positive report to a nonpositive one might affect the outcome, but if it does so then it changes the outcome from one in which the project is carried out to one in which the project is not carried out. Such a change makes the individual worse off (given that his valuation is positive). Thus for any reports of the other individuals, an individual with a positive valuation can do no better than report that valuation. A similar argument applies to an individual with a negative valuation.

Although the majority rule mechanism is strategy-proof, the condition it uses to determine whether the project is carried out has the disadvantage that it ignores the magnitudes of the individuals' valuations. If, for example, a few individuals would benefit hugely from the project and the remaining majority of

individuals would be made slightly worse off, then majority rule leads to the project's not being carried out.

An alternative mechanism, which takes into account the magnitudes of the individuals' valuations, carries out the project if the sum of the reported valuations is positive and, like majority rule, makes no monetary transfers. This mechanism, however, is not strategy-proof.

Example 17.2: Summing reports

Consider the direct mechanism in which the project is executed if and only if the sum of the individuals' reports is positive, and no monetary transfers are made. That is,

$$\delta(x^1,\ldots,x^n)=1 \text{ if and only if } \sum_{j\in N} x^j > 0$$

and $\tau^i(x^1,\ldots,x^n)=0$ for all $i\in N$ and for all profiles $(x^1,\ldots,x^n)$.

This mechanism is not strategy-proof. Consider an individual whose valuation is positive. If, when he reports his valuation, the sum of all reports is negative, so that the project is not carried out, then he is better off reporting a number high enough that the project is carried out.

We now describe a variant of the mechanism in this example that adds monetary transfers in such a way that the resulting mechanism is strategy-proof.

17.3 The Vickrey-Clarke-Groves mechanism

The Vickrey-Clarke-Groves (VCG) mechanism is a direct mechanism with transfers that executes the project if and only if the sum of the individuals' valuations is positive. The transfers in the mechanism are designed to make it strategy-proof: no individual benefits by reporting a number different from his true valuation. All transfers are nonpositive: under some circumstances an individual pays a penalty.

Suppose that, given the other individuals' reports, individual i's report is pivotal in the sense that given all the reports the project is executed, but in the absence of i's report it would not be. That is, the sum of all the reports is positive, but the sum of the reports of the individuals other than i is nonpositive. Then the monetary transfer for individual i in the VCG mechanism is equal to the sum of the other individuals' reports: i pays a penalty for causing the project to be executed when the other individuals' reports point to non-execution.

Now suppose that i's report is pivotal in the other direction: given all the reports the project is not executed, but in the absence of i's report it would be. (In particular, i's report is negative.) Then the monetary transfer for individual i is

the negative of the sum of the other individuals' reports. That is, i pays a penalty for causing the project not to be executed when the other individuals' reports point to execution.

Individual i pays a penalty only if his report makes a difference to the outcome. The penalty does not change when his report changes as long as the change does not affect the sign of the sum of the reports. If i's report is not pivotal in either sense, he pays no penalty.

Definition 17.5: VCG mechanism

For a public project problem $\langle N, D\rangle$, the *VCG mechanism* is the direct mechanism with transfers $(\delta, \tau^1, \ldots, \tau^n)$ defined by

$$\delta(x^1, \ldots, x^n) = 1 \text{ if and only if } \sum_{j \in N} x^j > 0$$

and

$$\tau^i(x^1, \ldots, x^n) = \begin{cases} \sum_{j \in N \setminus \{i\}} x^j & \text{if } \sum_{j \in N \setminus \{i\}} x^j \leq 0 \text{ and } \sum_{j \in N} x^j > 0 \\ -\sum_{j \in N \setminus \{i\}} x^j & \text{if } \sum_{j \in N \setminus \{i\}} x^j > 0 \text{ and } \sum_{j \in N} x^j \leq 0 \\ 0 & \text{otherwise.} \end{cases}$$

Notice that for some profiles of reports, the operator of the mechanism receives a positive amount of money from the individuals. It can be shown that for no strategy-proof direct mechanism do the transfers sum to zero for all possible profiles of reports (see Problem 1).

Here is a numerical example that illustrates the VCG mechanism.

Example 17.3: VCG mechanism

Consider the public project problem with $n = 5$. The following table shows the decision and transfers specified by the VCG mechanism for four profiles of reports.

x^1	x^2	x^3	x^4	x^5	δ	τ^1	τ^2	τ^3	τ^4	τ^5
5	−1	−1	−1	−1	1	−4	0	0	0	0
−5	1	1	1	1	0	−4	0	0	0	0
−7	1	1	3	4	1	0	0	0	−1	−2
−5	2	2	2	2	1	0	0	0	0	0

Proposition 17.1: VCG mechanism is strategy-proof

For any public project problem the VCG mechanism is strategy-proof.

Proof

Let $\langle N, D\rangle$ be a public project problem and let $(\delta, \tau^1, \ldots, \tau^n)$ be the VCG mechanism for this problem. Let $(v^i)_{i\in N}$ be a valuation profile.

Consider an individual $i \in N$ with $v^i > 0$, let $(x^1, \ldots, x^{i-1}, x^{i+1}, \ldots, x^n)$ be the reports of the other individuals, and let S be the sum of these other reports.

If $S > 0$ then if i reports v^i, the project is executed ($\delta(x^1, \ldots, x^n) = 1$) and his transfer is 0 ($\tau^i(x^1, \ldots, x^n) = 0$). Thus his utility is v^i. No outcome is better for him.

If $-v^i < S \leq 0$ then if i reports v^i or any other number greater than $-S$ the project is executed and his transfer is S, so that his utility is $S + v^i > 0$. If instead he reports a number at most $-S$ the project is not executed and his transfer is 0, so that his utility is only 0.

If $S \leq -v^i \leq 0$ then if i reports v^i or any other number at most $-S$ the project is not executed and his transfer is 0, so that his utility is 0. If instead he reports a number greater than $-S$ the project is executed and his transfer is S, so that his utility is $v^i + S \leq 0$.

Similar arguments apply if $v^i \leq 0$.

We conclude that for any reports of the other individuals and any valuation v^i, i's reporting v^i is not worse than his reporting any other number.

Discussion The VCG mechanism specifies that the project is carried out if and only if the sum of the individuals' valuations is positive. It is fairly simple, and relies only on the fact that no individual, regardless of his beliefs about the other individuals' reports, has any reason not to truthfully report his valuation. However, the following points diminish its appeal.

1. The outcome of the mechanism may require some individuals to make payments even if the project is not executed. For example, if the valuation profile is $(-5, 1, 1, 1, 1)$, then the project is not carried out and individual 1 makes a payment of 4. People may regard the requirement to pay money if the project is not carried out as unacceptable.

2. The mechanism is not very transparent; it takes time or experience to be persuaded that reporting one's valuation is indeed optimal independent of the other individuals' reports.

3. The payments are not distributed back to the individuals. If we change the mechanism so that the total payments collected are returned to the individuals then an individual's reporting his valuation is no longer necessarily

optimal for him regardless of the other individuals' reports. Consider, for example, a problem with two individuals, and assume that the total amount paid is distributed equally between the individuals. Suppose that $v^1 = 1$ and individual 2 reports 10. If individual 1 reports 1 the project is carried out and individual 1's utility is 1 (no payment is made). If individual 1 reports -8, however, the project is also carried out and individual 2 makes a payment of 8, half of which goes to individual 1, so that his utility is $1 + 4 = 5$. Thus individual 1 is better off reporting -8 than reporting his valuation of 1.

4. Using the sign of the sum of the valuations as the criterion for carrying out the project is not necessarily desirable, especially in a society in which the individuals differ widely in their wealths. Suppose that two individuals benefit slightly from the project, but due to their high wealth have valuations of 100 each. The other 99 individuals are hurt significantly by the project but are impoverished and have valuations of only -1. In this case the criterion requires that the project is carried out, even though it may seem unjust. The VCG mechanism not only requires that it is carried out but also that the wealthy make no payments. Their *willingness* to pay is enough to require the project to be carried out.

Problems

1. *Balanced budget.* A direct mechanism with transfers is *balanced* if, for all profiles of reports, the sum of the transfers is 0. A result that we do not prove states that there exists no strategy-proof balanced direct mechanism with transfers that carries out the project if and only if the sum of the valuations is positive. To illustrate this result, consider a public project problem with two individuals and assume that each individual's valuation is 4, -1, or -5.

   a. Find the outcome specified by the VCG mechanism for each report profile under the assumption that each individual is restricted to report only one of the three possible valuations, and verify that the mechanism is strategy-proof.

   b. Show that there exists no strategy-proof *balanced* direct mechanism with transfers for which the project is carried out only if the sum of the valuations is positive and the transfers are symmetric in the sense that the transfer for individual 1 when he reports x and individual 2 reports y is the same as the transfer for individual 2 when he reports x and individual 1 reports y.

2. *Vickrey auction.* One unit of a good is to be transferred to one of the individuals $1, \ldots, n$. Each individual i's valuation of the good is a nonnegative number v^i. Consider the following direct mechanism with transfers. Each individual reports a nonnegative number; the good is transferred to the individual, the *winner*, who reports the highest number. (In case of a tie, the good is transferred to the individual with the smallest index i among the individuals reporting the highest number.) The winner makes a payment equal to the highest of the other individuals' reports. (Thus if the reports are distinct, the winner's payment is the second highest report.) The winner's utility is his valuation minus his payment, and the utility of every other individual is 0. We can interpret this mechanism as a second-price auction, in which individuals submit bids and the good is transferred to the individual with the highest bid, who pays only the second highest bid (see Example 15.7).

   a. Show that the mechanism is strategy-proof.

   b. Explain why the direct mechanism with transfers that differs from the above only in that the winner makes a payment equal to his report is not strategy-proof.

3. *A project with a cost.* A group of n individuals has to decide whether to execute a project that costs C. If the project is executed, each individual pays $c = C/n$ to cover the costs. Individual i's utility from (α, t^i) is $\alpha(v^i - c) + t^i$ where α is 1 if the project is carried out and 0 if it is not, and t^i is a transfer.

   Design a VCG-like mechanism for this situation that is strategy-proof.

Notes

The idea behind the VCG mechanism is due to Clarke (1971) and Groves (1973), and has its origins in Vickrey (1961). Proposition 17.1 is established in Groves and Loeb (1975). The auction in Problem 2 was first studied by Vickrey (1961).