Contents

Lecture 10

Mechanism Design

A Sad Tale of Failed Design

Life in Our Town is simple. Folks here are interested in only three things. In Our Town, we still consider it indelicate to discuss one of them. So please assume that we are interested in only two things. These are hot dogs and the circus. There are only two kinds of people in town – the toads and the dudes. Toads don't care at all about the circus but always prefer more hot dogs to less. Dudes like both hot dogs and circus. As it happens, preferences of dudes can be represented by the utility function $U_D(X_i, Y) = X_i + 2\sqrt{Y}$ while toads' utility functions are simply $U_T(X_i, Y) = X_i$, where X_i is the amount of hot dogs that person i consumes and Y denotes the size of the circus. Each citizen, i, of Our Town has an initial endowment of wealth W_i which can be used either to buy hot dogs or to pay taxes. Tax revenue is used to pay for the circus. The bigger the circus, the more it costs. In fact, let us choose units of measurement for the size of the circus so that the cost of a circus of size Y is just \$$Y$. Let us also suppose that hot dogs cost \$1 each. There are N people in Our Town. Let us define an *allocation* to be a vector $(X_1, \cdots, X_N, Y)$ where X_i is the number of hot dogs consumed by person i and Y is the size of the circus. An allocation is *feasible* for the town if the total cost of hot dogs consumed plus the cost of the circus just equals total wealth of its citizens. The set of feasible allocations can therefore be denoted by

$$S = \{(X_1, \cdots, X_N, Y) | \sum_{i=1}^{N} X_i + Y = \sum_{i=1}^{N} W_i\}. \qquad (10.1)$$

A feasible allocation is said to be *Pareto optimal* if there is no other feasible allocation that is as good for everyone and better for someone. A classic result of Samuelson is that a necessary condition for Pareto optimality in a place like *Our Town* is that the sum of everyone's marginal rate of substitution of public for private goods must equal the marginal cost of public goods in terms of private goods. In our town the marginal cost of public goods is always one. Therefore the Samuelson condition takes the special form:

$$\sum_{i=1}^{N} \frac{\partial U_i}{\partial Y} \div \frac{\partial U_i}{\partial X_i} = 1. \tag{10.2}$$

Recalling the special form of utility functions assumed we see that for a toad $\frac{\partial U_i}{\partial Y} \div \frac{\partial U_i}{\partial X_i}$ is always zero. For a dude, we calculate $\frac{\partial U_i}{\partial Y} \div \frac{\partial U}{\partial X_i} = \frac{1}{\sqrt{Y}}$. Therefore in Our Town equation 10.2 takes the special form

$$N_D \frac{1}{\sqrt{Y}} = 1. \tag{10.3}$$

From 10.3 we see that the Pareto optimal amount of public goods for Our Town is

$$Y = N_D^2. \tag{10.4}$$

Our Town is a democracy. Everybody pays the same tax rate. We decide by majority vote how much circus to have. Of course toads always vote for no public goods, since they have to pay taxes but don't enjoy the circus. As it turns out, toads are in the minority in Our Town. Therefore dudes always out–vote the toads and get a positive amount of circus. (You might want to know why the toads haven't all moved to a town that has a majority of toads and no circus. The answer is that some of the necessary jobs in town can only be done by toads. For example, we need a banker, a mortician, some accountants, a realtor, a judge, and some school lunch monitors.)

How much circus would a dude like to have? Where Y is the amount of circus, his tax bill will be just $\frac{Y}{N}$. Therefore his after–tax wealth is just $W_i - \frac{Y}{N}$. Therefore he will be able to consume $X_i = W_i - \frac{Y}{N}$ hot dogs when the amount of circus is Y. His utility would then be

$$U_D(W_i - \frac{Y}{N}, Y) = W_i - \frac{Y}{N} + 2\sqrt{Y} \tag{10.5}$$

From 10.5 we see that

$$\frac{d}{dY} U_D(W_i - \frac{Y}{N}, Y) = \frac{1}{\sqrt{Y}} - \frac{1}{N}. \tag{10.6}$$

Therefore a dude's utility is an increasing function of Y for $Y < N^2$, a decreasing function of Y for $Y > N^2$ and is maximized at $Y = N^2$. Since there are more dudes than toads, it is clear that the only amount of circus that "wins" in majority voting is

$$Y = N^2. \tag{10.7}$$

For a long time the toads in Our Town have been grousing about high taxes and too much circus.dudes never paid much attention. The other day an economist visited us. (Claimed he wasn't a toad). He said the toads were right. He showed us equation 10.4 and pointed out that we have more than the Pareto efficient amount of public goods. He said he had just come from Their Town in the next county, where the problem was just the opposite. A majority of the people in Their Town (but not everyone) are toads. They have no circus at all.

This economist suggested that we try a different political system where we require unanimity instead of majority rule. But, since we have people with different tastes, we would have to set different tax rates for different people so as to get unanimity about quantities. He called this idea Lindahl equilibrium. In Our Town, the only way we could get the toads to agree to any positive amount of circus is if we don't tax them for the circus. Thendudes would have to pay all the taxes. Suppose that alldudes are taxed at the same rate. Then each dude would have a tax bill of $\frac{Y}{N_D}$. He could therefore consume $X_i = W_i - \frac{Y}{N_D}$ hot dogs and would have a utility of

$$U_D(W_i - \frac{Y}{N_D}, Y) = W_i - \frac{Y}{N_D} + 2\sqrt{Y} \tag{10.8}$$

This is maximized when $Y = N_D^2$. Therefore all dudes would choose the amount N_D^2 as their most preferred quantity of circus. Since toads pay no taxes and have no interest in the circus, this amount is as good as any other amount for them. Therefore the amount, N_D^2, receives unanimous approval. Therefore the Lindahl equilibrium is the allocation in which $Y = N_D^2$, $X_i = W_i$ if i is a toad and $X_i = W_i - \frac{Y}{N_D} = W_i - N_D$ if i is a dude.

The economist said that Lindahl equilibrium was both more equitable and more efficient than our old ways. The toads said he was right. The dudes were not so sure. A dude made the following calculations. Under the current system a dude has the utility:

$$W_i - \frac{N^2}{N} + 2\sqrt{N}^2 = W_i + N. \tag{10.9}$$

Under the Lindahl system a dude has the utility

$$W_i - \frac{N_D^2}{N_D} + 2\sqrt{N_D^2} = W_i + N_D. \tag{10.10}$$

Since $N > N_D$, moving to the Lindahl system is bad for dudes . The economist said that the dude had a point (though he was being a bit piggish). But the economist said that since we know that the current system is not Pareto optimal, it should be possible for the toads to bribe the dudes to move to Lindahl equilibrium. The economist pointed out that under the current system each toad has a utility of

$$X_i = W_i - \frac{N^2}{N} = W_i - N \tag{10.11}$$

while under the Lindahl system he would have no taxes so his utility would be

$$X_i = W_i. \tag{10.12}$$

We can see from expressions (9) and (10) that a dude could be bribed to accept the Lindahl system if he was given $N - N_D = N_T$ hot dogs. Since there are N_D dudes , it would take $N_D N_T$ hot dogs to bribe all of the dudes to accept the Lindahl system. Therefore if each toad gave up N_D hot dogs to bribe the dudes , there would be just enough hot dogs to do so. If this is done, each toad would have a utility of

$$X_i = W_i - N_D. \tag{10.13}$$

Equation expresses the utility of each dude in the Lindahl system without bribes. With bribes of N_T for each dude, the utility of each dude would be

$$W_i + N_D + N_T = W_i + N \tag{10.14}$$

which is the same as his utility under the current system. Since Expression 10.13 is greater than Expression 10.11 and since Expression 10.14 equals Expression 10.9, we see that moving to Lindahl equilibrium with this system of bribes benefits all toads and leaves all dudes as well as before. If we made the bribes slightly larger, everyone would be better off than in the current system.

The dudes and the toads were all impressed by this argument. The bribes were paid, and the entire community agreed to switch to the Lindahl system. There was one small hitch. You can't always tell by looking, whether a person is a toad or a dude. To solve this problem, the mayor asked everyone to come down to the town hall and answer the simple question:

"Are you a dude?"

To his amazement almost everyone sauntered in and said

"Yeah, man."

How did this happen? Toads being a thoughtful lot, each toad asked itself, would I be better off pretending to be a dude? If everybody else is telling the truth, then if I confess to being a toad, my share of the cost of bribing the dudes to accept the Lindahl system will be N_D. Since my Lindahl tax will be 0 and I don't care at all about the circus, my utility will be $W_i - N_D$. But what if I claim to be a dude? Then the number of dudes registered in the town hall will be $N_D + 1$ and the number of registered toads will be $N_T + 1$. In this case, my Lindahl tax will be $(N_D + 1)^2 / (N_D + 1) = N_D + 1$ and I will receive a bribe from the other toads for accepting the Lindahl system. The bribe that I get will also be equal to $N_D + 1$. Since my bribe is equal to my Lindahl tax, I will be able to consume W_i hot dogs and my utility will be W_i. So if the others all tell the truth, I am best off claiming to be a dude.[1]

The mayor calculated and provided the Lindahl equilibrium amount of circus and the distribution of taxes, given the reported number of dudes and toads. Since everyone claimed to be a dude, the Lindahl quantity of circus was N^2, just as it had been before the reforms were introduced. Also, just as before, each individual in town paid a tax of $N^2/N = N$. Curiously enough, even though the folks in Our Town took the economists' advice to heart and acted on it, the outcome was no different than it had been before the economist rode into town.[2]

As a result of this sad experience, dudes in Our Town are inclined to look at economists (and at each other) with suspicion. True toads, of course, are

[1] One or two of the deeper-thinking toads thought further along these lines. If it pays me to claim to be a dude, then perhaps some of the other toads will also notice that it pays them to claim to be dudes. And for that matter, how do I know that the real dudes will want to admit to be dudes? Such a toad would think as follows: Suppose that M_D residents claim to be dudes and $N - M_D = M_T$ claim to be toads. Then if I admit to being a toad, I will pay M_D as my share of the cost of bribing the alleged dudes to accept the Lindahl mechanism so my utility will be $W_i - M_D$. If I claim to be a dude, I will have Lindahl taxes of $M_D^2/M_D = M_D$ and I will receive a bribe of M_D from the alleged toads for accepting the Lindahl system. Since my bribe is equal to my Lindahl tax, my utility if I claim to be a dude will be W_i, which is greater than my utility if I confess to being a toad.

[2] A tactless political scientist might remark that this is one of the rare occasions when economists' advice did no harm.

pleased and amused with the outcome.

It is time, I think, to draw the curtain on the sordid situation in Our Town, while we seek aid from some more general analysis. So far, we have learned the following lessons which apply not only in Our Town but quite generally.

1. For an arbitrary distribution of taxes, majority voting will not in general lead to a Pareto optimal supply of public goods.

2. Lindahl equilibrium is Pareto optimal. However imposition of a Lindahl equilibrium requires the central authority to know individual preferences.

3. If people are asked to state their preferences, knowing that their statements will be used to calculate a Lindahl equilibrium that will then be imposed, the situation where everyone tells the truth is *not* a best response (Nash) equilibrium.

The difficulty in item (3) is often called the "preference revelation problem". It is representative of a fascinating class of problems of the firm. "How do you get someone else to tell you the truth about something that only he knows?"

Eliciting the Truth via the Pivotal Mechanism

A philosopher who dabbles in economics, Alan Gibbard [3], and an economist who dabbles in philosophy, Mark Satterthwaite [4], independently showed that in general it is not possible to design such mechanisms.

There are, however, some interesting special cases where theory suggests that a "nearly" Pareto optimal outcome can be implemented by a mechanism in which it is a weakly dominant strategy to tell the truth.

A strategy is defined to be *weakly dominant* if one cannot do better than to use this strategy, no matter what anybody else does. A strategy is defined to be *strictly dominant* if it is always better to use this strategy rather than any other, no matter what others are doing. You can never gain by deviating from a weakly dominant strategy, but possibly you might not lose. If you deviate from a strictly dominant strategy, you will surely be worse off. (Frequently, in the literature, the term *dominant strategy* is applied to either case.)

A Second-price Sealed Bid Auction

One method that selects a Pareto efficient outcome is found by applying William Vickrey's idea of a *second-price sealed bid auction* to public decision making. Recall the way the second-price sealed bid auction works. There are n people and one object to the allocated among them. Let V_i be the maximum amount that person i would be willing to pay for the object. Pareto efficient allocations would have the object go to the person with the greatest willingness to pay. If a sealed-bid auction were held, with the object going to the highest bidder at his bid price, it would not be wise for anyone to bid his true valuation. In the second-price sealed bid auction, the object is awarded to the highest bidder who pays the bid made by the second highest bidder. With this system, it turns out that bidding ones true valuation is the best thing to do *no matter what other people bid.* A strategy that is best no matter what others do is known as a *dominant strategy.* A social outcome where everyone is using a weakly dominant strategy is called a *dominant strategy equilibrium.* In Vickrey's auction, the outcome where everyone bids his true valuation and the object goes to the person with the highest valuation at price equal to the second highest valuation is a dominant strategy equilibrium. Lets see why this is so. Suppose that you bid more than your true evaluation. If your bid is not the highest bid, you are no better (or worse) off than if you had told the truth. If your bid is the highest bid, then there are two possible cases. If your true valuation would also have been the highest bid, then you are no better (or worse) off than if you had bid the truth. If your true valuation is lower than the second highest bid, then you get the object but you must pay more than it is worth to you. You would have been better off bidding the truth and not getting the object. Thus we see that you can not gain but you can lose by overbidding. You should be able to construct a similar argument to show that you can not gain and may lose by underbidding. Therefore, bidding the truth is a dominant strategy.

The Pivotal Mechanism

The idea of Vickrey's auction can be extended to other kinds of discrete choices. Of particular interest are yes-or-no choices on public issues, where the outcome has negligible direct financial cost. Political issues of this type include whether to allow the sale of handguns, legalize abortion, or to allow the sale of marijuana. Similar issues arise for smaller groups. For example, a class may want to decide whether to have its midterm exam on Tuesday or Thursday.

One possible decision mechanism is simple majority vote. The weakness of this mechanism is that it may not be Pareto optimal. The minority may be intensely concerned, while members of the majority each care very little. In this case it might be possible to find a Pareto superior outcome which reverses the majority voting result since the minority cares enough to buy off the majority.

Suppose that a community currently permits possession of handguns, but is considering whether to ban them. Citizen i's utility depends on the amount of private goods X_i that she consumes and whether the ban is passed. Her utility function is $U_i(X_i, 1)$ if handguns are banned and $U_i(X_i, 0)$ if handguns are not banned. Let i's willingness to pay for the ban be the quantity of private goods V_i such that

$$U_i(X_i - V_i, 1) = U_i(X_i, 0). \tag{10.15}$$

Then $V_i > 0$ for those who favor the ban and $V_i < 0$ for those who oppose it. It is not difficult to show that with appropriate monetary side-payments, everyone in the community could be better off with a ban on handguns if and only if $\sum V_i > 0$.

If we just asked people to state their willingnesses to pay and then decided the issue by the sum of these stated willingnesses, individuals would have an incentive to overstate the intensity of their preferences. We need a more subtle device. The following scheme, known as the *pivotal mechanism*, is designed to elicit true statements of the V_i's and to make the appropriate social decision. Each i is asked to state his willingness to pay to have the ban on handguns. Where M_i is i's response, the ban will be passed if and only if $\sum_i M_i > 0$. Answers will be truthful if $M_i = V_i$ for all i. To motivate truthful answers, taxes are assessed as follows. Person i is said to be *pivotal* if and only if her answer changes the outcome. That is, if and only if the sign of $\sum_{j \neq i} M_j$ is the opposite of the sign of $\sum_j M_j$. If i is not pivotal, she pays no tax. If person i is pivotal, then she pays a tax equal to the amount $\sum_{j \neq 1} M_j$. Any revenue from this scheme is thrown away. Using exactly the same kind of reasoning that we used for the Vickrey auction, we can show that telling the truth is a weakly dominant strategy for every player. If everyone tells the truth, the ban will be passed if and only if it would be possible to arrange side-payments from the gainers to the losers so that everyone is better off than in the status quo.

Weaknesses of the Pivotal Mechanism

While the pivotal mechanism is very interesting, at least on theoretical grounds, it has some troublesome weaknesses.

1. In general, the outcome is not efficient because any tax revenue that is collected must be thrown away.

   As a practical matter, this problem may not be very severe. First, because in a reasonably large community, it is very likely that nobody or a very small number of people will actually be pivotal, so the expected revenue is small. (I wish I had a decisive reference on this matter, or a good proof, but at the moment I do not.) Secondly, if more than one organization used the pivotal mechanism, each could agree to give any revenue it collected to the other. This would not influence individual incentives, since nobody can affect the outcome in a community other than his own.

2. Although the pivotal mechanism approves a change only if the gainers could potentially compensate the losers, the mechanism does not implement this redistribution.

   One might defend the pivotal mechanism against this charge by arguing that if this mechanism is repeatedly in a community to decide issues, then the "law of large numbers" suggests with high probability everyone will be better off using a mechanism that maximizes the sum of willingnesses to pay.

   But even if the mechanism is used many times, it is likely to more favorable to the rich and less favorable to the poor than majority voting. If one thinks that the distribution of political rights should be more equal than that of income, this is a serious consideration.

3. Two or more colluding participants can "cheat" the pivotal mechanism by sending false signals.

   Suppose that two people who agree on the preferred outcome agree to each state a number that is certain to be much larger in absolute value than the sum of everyone else's willingness to pay. They will get their way and neither will be pivotal, so neither has to pay anything.

4. People may not be convinced that reporting their true valuations is a weakly dominant strategy and hence may not report the truth.

In a set of experiments conducted by Attiyeh, Franciosi, and Isaac [1] most participants do not respond truthfully and the mechanism does not choose the "efficient" outcome more often than does simple majority voting. Related experiments are discussed in a recent survey by Chen and Ledyard [2].

Although truth-telling is a *weakly dominant strategy* in the pivotal mechanism, it is not *strictly dominant.* That is, you can not gain by lying, but you might not lose anything by doing so. If participants do not believe that truth telling is a weakly dominant strategy, it is not easy for them to learn this from experience.

Exercises

10.1 Explain why it is that in the pivotal mechanism, it is a weakly dominant strategy to tell the truth. Write your explanation in such a way that it is convincing for an intelligent person who has not studied game theory.

10.2 Prove the assertion in the text that a Pareto improvement could be achieved by passing the ban and making appropriate side-payments if $\sum V_i > 0$, and could not be achieved if $\sum V_i < 0$.

Bibliography

[1] Greg Attiyeh, Robert Franciosi, and R. Mark Isaac. Experiments with the pivot process for providing public goods. *Public Choice*, 102(1):95–114, January 2000.

[2] Yan Chen and John Ledyard. Mechanism design experiments. to appear in the New Palgrave, 2008, October 2006.

[3] Alan Gibbard. Manipulation of voting schemes: A general result. *Econometrica*, 41(4):587–601, July 1973.

[4] Mark Satterthwaite. Strategy-proofness and arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *Journal of Economic Theory*, 10(2):187–217, April 1976.

Bibliography

[1] Greg Attiyeh, Robert Franciosi, and R. Mark Isaac. Experiments with the pivot process for providing public goods. *Public Choice*, 102(1):95–114, January 2000.

[2] Yan Chen and John Ledyard. Mechanism design experiments. to appear in the New Palgrave, 2008, October 2006.

[3] Alan Gibbard. Manipulation of voting schemes: A general result. *Econometrica*, 41(4):587–601, July 1973.

[4] Mark Satterthwaite. Strategy-proofness and arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *Journal of Economic Theory*, 10(2):187–217, April 1976.