0.0.34 Roger Myerson and Mark Satterthwaite, "Efficient Mechanisms for Bilateral Trading," *Journal of Economic Theory*, Vol 29, p. 265-281 (1983).

The paper concerns the possible trade of a single, indivisible item between a buyer and a seller. A key aspect of the model is that each trader privately knows the value to him of trading. We will focus on the following result: it is impossible for trading to be perfectly efficient in any *plausible* bargaining mechanism. Defining "plausible" is part of the formal analysis of the paper.

0.0.35 The Model

The buyer's type is his value $v \in [\underline{v}, \overline{v}]$ and the seller's type is his cost $c \in [\underline{c}, \overline{c}]$. The buyer's value is drawn from a distribution $G(\cdot)$ whose density $g(\cdot)$ is nonzero on $[\underline{v}, \overline{v}]$, while the seller's cost is drawn from the distribution $F(\cdot)$ whose density $f(\cdot)$ is nonzero on $[\underline{c}, \overline{c}]$. A trader observes his own type but not the type of the other trader. The buyer's ex post utility from acquiring the item is

$$v - p,$$

where p is any payment that he makes. The buyer's utility is $-p$ when he makes the payment p and does not receive the item. Similarly, the buyer's ex post utility from selling the item is

$$p - c,$$

where p is any payment that he receives. The seller's utility is p when he receives the payment p and does not give up the item the item.

A **revelation mechanism** consists of an allocation rule $q(v, c)$ and $p(v, c)$, where $q(v, c)$ specifies the probability that the item is transferred and $p(v, c)$ the payment from buyer to seller when v and c are reported. Notice that $p(v, c)$ can be nonzero (money can change hands) even when the item is not transferred. The *efficient allocation rule* $q(v, c)$ requires that

$$q(v, c) = \begin{cases} 1 \text{ if } v > c \\ 0 \text{ if } v < c \end{cases}.$$

We'll focus on it in this discussion.We define the interim probabilities and expected payments in the case of the efficient allocation rule as follows:

$$Q_b(v) = \int_{\underline{c}}^{\overline{c}} q(v, c) f(c) dc,$$

$$P_b(v) = \int_{\underline{c}}^{\overline{c}} p(v, c) f(c) dc,$$

$$\Pi_b(v, v^*) = v Q_b(v^*) - P_b(v^*)$$

$$\Pi_b(v) = \Pi_b(v, v),$$

$$Q_s(c) = \int_{\underline{v}}^{\overline{v}} q(v, c) g(v) dv,$$

$$P_s(c) = \int_{\underline{v}}^{\overline{v}} p(v,c)g(v)dv,$$
$$\Pi_s(c,c^*) = P_s(c^*) - cQ_s(c^*),$$
$$\Pi_s(c) = \Pi_s(c,c).$$

We impose the following constraints on revelation mechanisms:

- **incentive compatibility (IC)**: for all $v, v^* \in [\underline{v}, \overline{v}]$,
$$\Pi_b(v) \geq \Pi_b(v,v^*),$$
and for all $c, c^* \in [\underline{c}, \overline{c}]$,
$$\Pi_s(c) \geq \Pi_s(c,c^*).$$
- **(interim) individual rationality (IR)**: for all $v \in [\underline{v}, \overline{v}]$,
$$\Pi_b(v) \geq 0,$$
and for all $c \in [\underline{c}, \overline{c}]$,
$$\Pi_s(c) \geq 0.$$
- We've also implicitly included the constraint of **budget balance** by assuming that the price $p(v,c)$ paid by a buyer is received by the seller (i.e., there is neither "leakage" of money from the pair of bargainers nor any external subsidy, in any state (v,c)).

Through the Revelation Principle, all possible outcomes of individually rational, Bayesian-Nash equilibria of all possible bargaining procedures that are ex post budget-balanced are addressed by studying revelation mechanisms (q,p) that satisfy IC and IR. This will be discussed further below.

0.0.36 Main Result

The following result will be proven:

Theorem 25 *Assume that*
$$(\underline{v}, \overline{v}) \cap (\underline{c}, \overline{c}) \neq \varnothing,$$
i.e., it is not certain ex ante that the item either should be traded or should not be traded. For the efficient allocation rule $q(v,c)$, there does not exist a pricing rule $p(v,c)$ so that IC and IR are satisfied.

This result is particularly meaningful for economic thought. Classical price theory trivializes the process of bargaining and simply asserts that traders achieve all possible gains from trading without modeling the process by which these gains are achieved. The Myerson-Satterthwaite result is quite different, for it asserts that trading cannot always be efficient. Sometimes, potential gains from trading between two traders will not be achieved, and this loss is unavoidable. The lost gains from trade may appear as costly delay before agreements are reached (as in labor strikes).

Contrast this result with the "Coase Theorem" or "Coase Conjecture" of Ronald Coase. Coase argued that the problem of externalities is fundamentally one of confusion over property rights. The role of the legal system is then to clarify property rights between the party creating the externality and the party that is either harmed or benefits from the externality. Once property rights are clarified, the two parties then can then (efficiently) negotiate between themselves to regulate the externality. The Coase Conjecture is that this negotiation is efficient and therefore solves the problem of the externality. In contrast, the Myerson-Satterthwaite result suggests that the externality may not be efficiently regulated by the two parties even if property rights have been clarified.

Example 80 *Coase proposed the example of a rancher whose cattle trample the crops of a farmer on their way from one pasture to another. According to Coase, the essential problem is whether or not the rancher is entitled to move his cattle across the farmer's fields. The role of the government is to clarify this issue, and how it is decided really is not important to the efficiency of the outcome. If, for instance, the rancher has the right to move his cattle across the fields, and if both the rancher and the farmer know that the rancher has this right, then the farmer may compensate the rancher so as to limit the damage. If the rancher does not have this right, then he can negotiate access with the farmer in return for a payment From the Myerson-Satterthwaite perspective, Coase trivializes the process subsequent to the clarification of property rights in which the two parties negotiate. They assert that the outcome of negotiation may fail to be an efficient resolution of the problem of externalities.*

0.0.37 What does the VCG mechanism have to say about this bargaining model?

The mechanism operates as follows in this case. Given reported values and costs v^*, c^*, the item is traded iff $v^* \geq c^*$, in which case the buyer receives the transfer $-c^* + t_b(c^*)$ (or makes the payment $c - t_b(c^*)$) and the seller receives $v^* + t_s(v^*)$. The item is not traded if $v^* < c^*$, in which case the buyer receives $t_b(c^*)$ and the seller receives $t_s(v^*)$.

This mechanism is ex post efficient and it satisfies IC. To satisfy ex post budget balance, we would need

$$t_b(c) + t_s(v) = 0 \textit{ if } c < v$$

and

$$-c + t_b(c) + v + t_s(v) = 0 \textit{ if } v \geq c.$$

The top equation implies that $t_b(c), t_s(v)$ are constants and $t_b = -t_s$, which contradicts the second equation. No member of the family of VCG mechanisms can therefore satisfy ex post budget balance.

We can choose $t_b(c)$ and $t_s(v)$ so that the mechanism has a deficit/surplus of zero ex ante. The *basic* VCG mechanism charges the buyer a lower price c than the payment v to the buyer when trade occurs ($v \geq c$), and each trader's expected gains from trade equals the entire gains from trade,

$$\Gamma = \int_{\underline{c}}^{\overline{c}} \int_{\underline{v}}^{\overline{v}} (v - c)\, q(v, c) g(v) f(c) dv dc.$$

Note that we'll use Γ in what follows to denote the expected gains from trade in the efficient allocation rule. The deficit of the basic VCG mechanism therefore equals Γ. Keeping in mind that the mechanism "collects" $-(t_b + t_s)$ from the traders, we simply need to pick constants t_b, t_s so that

$$-(t_b + t_s) = \Gamma.$$

A problem with such individualized transfers, however, is that at least one of these two constants must be negative, which means that the corresponding VCG mechanism may not be individually rational. The point, however, is that the VCG mechanism insures efficient trade if one is willing to forego budget balance and individual rationality. We'll use the VCG mechanism in our analysis of the Myerson-Satterthwaite model below.

0.0.38 Digression: The Revelation Principle

Why do we focus exclusively on bargaining in incentive compatible revelation mechanisms? Lying is common in bargaining; why do we insist that traders be given the incentive to be honest? Could the efficient allocation rule be implemented in some bargaining game in which a trader's strategy may be more complex than a simple report of his value/cost, or one in which he has the incentive to not reveal his true value/cost? The answer to these questions lies in the *Revelation Principle*. It states that any possible equilibrium of any possible bargaining game can be replicated or implemented as the outcome of some incentive compatible revelation game. By studying incentive compatible outcomes of revelation games, we effectively study all possible equilibrium outcomes of all possible bargaining games.

Let $\mathcal{L}$ denote the set of all possible rankings of the elements in a set of alternatives A. Notation for a game: agent i's *strategy set* is S_i, $\sigma_i : \mathcal{L} \to S_i$ denotes a *strategy* of player i (i.e., a choice based upon his preferences over the alternatives A), and

$$\tau : \prod_{i=1}^{N} S_i \to A$$

denotes the *outcome function* of the game (i.e., how the game determines an alternative based upon the strategic choices of the players). A *game* or *mechanism* is denoted

$$\left(\prod_{i=1}^{N} S_i, \tau \right).$$

Letting σ:$\mathcal{L}^N \to \prod_{i=1}^{N} S_i$ be defined by

$$\sigma = (\sigma_1, \ldots, \sigma_N),$$

the social choice function f is *implemented* by the strategy profile $(\sigma_i)_{1 \leq i \leq N}$ in the game $(\prod S_i, \tau)$ if

$$\tau \circ \sigma = f,$$

i.e., f results when the agents employ the strategies $(\sigma_i)_{1 \leq i \leq N}$.

A *revelation game* or mechanism is a game in which each $S_i = \mathcal{L}$ (i.e., each agent is given the opportunity to report his complete ranking of the alternatives in A).

Theorem 26 (Revelation Principle for Dominant Strategies) *If $(\sigma_i)_{1 \leq i \leq N}$ is a dominant strategy equilibrium in the game $(\prod S_i, \tau)$, then honest revelation by each agent is a dominant strategy equilibrium in the game $\left(\mathcal{L}^N, \tau \circ \sigma\right)$.*

Proof. We need to show that for each agent i and every $(L_1, \ldots, L_i, \ldots, L_N)$, and every $L_i^{'}$,

$$\tau \circ \sigma(L_1, \ldots, L_i, \ldots, L_N) \geq_{L_i} \tau \circ \sigma(L_1, \ldots, L_i^{'}, \ldots, L_N)$$

We have

$$\begin{aligned} \tau \circ \sigma(L_1, \ldots, L_i, \ldots, L_N) &= \tau(\sigma_1 (L_1), \ldots, \sigma_i (L_i), \ldots, \sigma_N (L_N)) \\ &\geq {}_{L_i}\tau(\sigma_1 (L_1), \ldots, \sigma_i \left(L_i^{'}\right), \ldots, \sigma_N (L_N)) \\ &= \tau \circ \sigma(L_1, \ldots, L_i^{'}, \ldots, L_N), \end{aligned}$$

where the inequality is true because $(\sigma_i)_{1 \leq i \leq N}$ is a dominant strategy equilibrium. ■

A verbal explanation of the Revelation Principle is as follows. Suppose we are given the dominant strategy equilibrium $(\sigma_i)_{1 \leq i \leq N}$ in the game $(\prod S_i, \tau)$. Imagine that we ask each agent i, "Report to me (as a nonstrategic and honest "operator" of the game) your preferences and I will carry out for you in the game $(\prod S_i, \tau)$ the action you would take according to the strategy σ_i." The claim is that it is a dominant strategy for agent i to be honest in his report, for if he had some reason to lie for some profile of reported preferences of the other players, then he would also have reason to not use σ_i in the game $(\prod S_i, \tau)$.

The Revelation Principle also holds for Bayesian-Nash equilibrium, with honest reporting a Bayesian-Nash equilibrium in the revelation game.

0.0.38.1 The Meaning of the Revelation Principle

The implication of the revelation principle is that we can examine *all* social choice functions that result from dominant strategy equilibria of *all* possible games simply by considering those that result from honest revelation being a dominant strategy in a revelation mechanism. In other words, the set of all strategy-proof social choice functions consists of all social choice functions that can be implemented as dominant strategy equilibria of arbitrary games. This is important because we may not want to be obsessed with truth-telling as an end in itself, or solely with revelation games.

Example 81 *Consider the following two-player game of incomplete information:*

1/2	*L*	*R*
T	$2\theta_1, 3\theta_2$	$1, 1$
B	$1, 0$	$0, 1$

It is common knowledge among the two players that each player i's type θ_i is independently drawn from the uniform distribution on $[0, 1]$.

- *Derive a pure strategy Bayesian-Nash equilibrium in this game.*
- *Illustrate the revelation principle by defining a revelation game that results in the outcome derived for your answer in a) and has the property that truthful revelation is a Bayesian-Nash equilibrium.*

We first note that player 1 has a dominant strategy to choose T for each value of his type $\theta_1 > 1/2$ and player 2 has a dominant strategy to choose R for each value of his type

$\theta_2 < 1/3$. We therefore conjecture the following form of equilibrium strategies:

$$\text{player 1: } \begin{cases} T \text{ if } \theta_1 > \theta_1^* \\ B \text{ if } \theta_1 \leq \theta_1^* \end{cases},$$

$$\text{player 2: } \begin{cases} L \text{ if } \theta_2 \geq \theta_2^* \\ R \text{ if } \theta_2 < \theta_2^* \end{cases}.$$

Solving for equilibrium requires solving for the constants θ_1^* and θ_2^*. These are the "critical" values of each player's type at which he is indifferent between each of his two pure strategies. Notice that player 1 plays T with probability $1 - \theta_1^*$ and B with probability θ_1^*, while player 2 plays L with probability $1 - \theta_2^*$ and R with probability θ_2^*. We have the equation

$$2\theta_1^* \cdot (1 - \theta_2^*) + 1 \cdot \theta_2^* = 1 \cdot (1 - \theta_2^*) + 0 \cdot \theta_2^*$$

for player 1, expressing his indifference between B and T when θ_1^* is his type, and the equation

$$3\theta_2^* \cdot (1 - \theta_1^*) + 0 \cdot \theta_1^* = 1 \cdot (1 - \theta_1^*) + 1 \cdot \theta_1^*$$

for player 2, expressing his indifference between L and R when θ_2^* is his type. Player 1's equation reduces to

$$\theta_1^* (1 - \theta_2^*) + \theta_2^* = \frac{1}{2} \Leftrightarrow \theta_2^* (1 - \theta_1^*) + \theta_1^* = \frac{1}{2}$$

and player 2's equation reduces to

$$\theta_2^* (1 - \theta_1^*) = \frac{1}{3}.$$

Substituting the second equation into the first produces

$$\frac{1}{3} + \theta_1^* = \frac{1}{2} \Rightarrow \theta_1^* = \frac{1}{6}.$$

We can then solve for θ_2^* as

$$\theta_2^* \left(\frac{5}{6}\right) = \frac{1}{3} \Rightarrow \theta_2^* = \frac{2}{5}.$$

Given the strategy for player 2 determined by $\theta_2^* = 2/5$, the expected difference for player 1 between choosing T and choosing B when θ_1 is his type equals

$$\left[2\theta_1 \cdot \frac{3}{5} + 1 \cdot \frac{2}{5}\right] - \left[1 \cdot \frac{3}{5} + 0 \cdot \frac{2}{5}\right]$$
$$= \frac{1}{5} [6\theta_1 + 2 - 3]$$
$$= \frac{1}{5} [6\theta_1 - 1],$$

which changes from negative to positive at $\theta_1 = 1/6 = \theta_1^*$. This supports his use of the strategy that we have derived for him. Similarly, given the use of this strategy for player 1, the difference in player 2's expected payoff between L and R when his type is θ_2 equals

$$\left[3\theta_2 \cdot \frac{5}{6} + 0 \cdot \frac{1}{6}\right] - \left[1 \cdot \frac{5}{6} + 1 \cdot \frac{1}{6}\right]$$
$$= \left[\theta_2 \cdot \frac{5}{2} - 1\right].$$

This changes from negative to positive at $\theta_2^* = 2/5$, which supports player 2's use of the *strategy that we derived above.*

Our equilibrium is therefore

$$\textit{player 1:}\begin{cases} T \textit{ if } \theta_1 > \frac{1}{6} \\ B \textit{ if } \theta_1 \leq \frac{1}{6} \end{cases},$$

$$\textit{player 2:}\begin{cases} L \textit{ if } \theta_2 \geq \frac{2}{5} \\ R \textit{ if } \theta_2 < \frac{2}{5} \end{cases}.$$

Revelation game: Let θ_1^* and θ_2^* now denote the reports of the two agents. The revelation game is defined as follows

$\theta_1^* \leq \frac{1}{6}, \theta_2^* < \frac{2}{5}$ *: B,R (0,1)*
$\theta_1^* > \frac{1}{6}, \theta_2^* < \frac{2}{5}$ *: T,R (1,1)*
$\theta_1^* \leq \frac{1}{6}, \theta_2^* \geq \frac{2}{5}$ *: B,L (1,0)*
$\theta_1^* > \frac{1}{6}, \theta_2^* \geq \frac{2}{5}$ *: T,L* $(2\theta_1, 3\theta_2)$.

0.0.38.2 A Lemma

Lemma 27 *For the efficient allocation rule $q(v,c)$, suppose the pricing rule $p(v,c)$ satisfies IC. Then:*

1.
$$\begin{aligned} \Pi_b(v) &= \Pi_b(\underline{v}) + \int_{\underline{v}}^{v} Q_b(y)\, dy, \qquad (21) \\ \Pi_s(c) &= \Pi_s(\overline{c}) + \int_{c}^{\overline{c}} Q_s(x)\, dx, \end{aligned}$$

where x and y are dummy variables.

2. *If IR is also satisfied by $q(v,c)$ and $p(v,c)$, then $q(v,c)$ satisfies the equation*

$$\Gamma = \Pi_b(\underline{v}) + \int_{\underline{v}}^{\overline{v}} \int_{\underline{v}}^{v} Q_b(y)\, g(v) dy dv + \Pi_s(\overline{c}) + \int_{\underline{c}}^{\overline{c}} \int_{c}^{\overline{c}} Q_s(x)\, f(c) dx dc. \qquad (22)$$

The x and the y in (22) are dummy variables. Notice that from 2. it follows that IR is satisfied by an IC mechanism if and only if

$$\Pi_b(\underline{v}), \Pi_s(\overline{c}) \geq 0.$$

We can think of $\underline{v}$ as the least favorable value for the buyer and $\overline{c}$ as the least favorable cost for the seller. Given IC, IR thus binds as a constraint at the "worst-off" types of the two traders.

Notice also that the formulas in (21) are derived directly from the constraint of incentive compatibility. IC is quite restrictive as a constraint; I sometimes refer to it as a double continuum of constraints, because each pair of a true value and a reported value defines its own inequality. IC is so restrictive as a constraint that *it determines each trader's profit function $\Pi_b(v)$, $\Pi_s(\overline{c})$ up to a constant ($\Pi_b(\underline{v})$ or $\Pi_s(\overline{c})$).*

Proof. We'll establish 1. for $Q_s(c)$. A similar argument applies to establish the results for $Q_b(v)$. For $c, c^* \in [\underline{c}, \overline{c}]$, IC together with the Envelope Theorem implies

$$\frac{d\Pi_s}{dc}(c) = \frac{\partial \Pi_s}{\partial c}(c, c^*)\,|_{c^*=c} = -Q_s(c)\,.$$

Taking antiderivatives and solving for the constant by which they may differ implies

$$\Pi_s(c) = \Pi_s(\overline{c}) + \int_c^{\overline{c}} Q_s(x)\,dx$$

as in (21).

Inequality (22) starts from the equality

$$\Gamma = \int_{\underline{v}}^{\overline{v}} \Pi_b(v)\,g(v)dv + \int_{\underline{c}}^{\overline{c}} \Pi_s(c)\,f(c)dc,$$

i.e., what the traders end up with from bargaining (the right hand side) equals the expected gains from trade. Substitute for $\Pi_b(v)$ and $\Pi_s(c)$ using (21) and move the constants $\Pi_b(\underline{v})$, $\Pi_s(\overline{c})$ to the right hand side. IR implies that $\Pi_b(\underline{v}) + \Pi_s(\overline{c}) \geq 0$, which produces (22). ■