The following paper was originally published in the
Proceedings of the First USENIX Workshop on Electronic Commerce
New York, New York, July 1995.

Economic Mechanism Design for Computerized Agents

Hal R. Varian
School of Information Management and Systems
University of California, Berkeley, CA

Economic Mechanism Design for Computerized Agents

Hal R. Varian
School of Information Management and Systems
University of California
Berkeley, CA 94720
hal@sims.berkeley.edu

Abstract

The field of *economic mechanism design* has been an active area of research in economics for at least 20 years. This field uses the tools of economics and game theory to design "rules of interaction" for economic transactions that will, in principle, yield some desired outcome. In this paper I provide an overview of this subject for an audience interested in applications to electronic commerce and discuss some special problems that arise in this context.

1 Mechanism design

As an example of mechanism design in action, let us consider the case of designing an auction to award an item to one of n individuals. Each individual i has a "maximum willingness to pay" or "value" for the item that we denote by v_i. We assume that this value is *private information* known only by person i. Our goal is to design an auction that will award the item to the person with the highest value.

*This work was supported by the National Science Foundation Grant SES-93-20481. I wish to thank Jeffrey K. MacKie-Mason for his helpful comments on an earlier draft.

The most obvious way to do this is to use a standard *English auction*. In this game, the auctioneer continuously raises the price of the good. Bidders who are unwilling to pay the current price drop out until only one bidder is left. It is not hard to see that this remaining bidder must be the person with the highest value. However, it is important to observe that the price that he pays for the good will be the willingness to pay of the person with the *second* highest value (plus, perhaps, a tiny amount to break the tie).

This sort of auction is fine when communication costs are low and iteration is cheap. But what if communication costs are high? For example, suppose that one wants to conduct an auction that is distributed over space and/or time. Is there a way to achieve the result of the English auction without iteration? A standard form of one-shot auction is the *sealed bid auction*. In this game, each player submits a sealed bid. The bids are opened and the item is awarded to the person with the highest bid. That person in turn pays the price he bid for the good.

This auction avoids iteration but it will not in general achieve the desired objective of awarding the item to the bidder with the highest value. Suppose that bidder 1 has value of 1 and bidder 2 has value of 2. However, bidder 2 mistakenly *believes* that bidder 1's value is 1/2. Bidder 2 therefore bids $1/2 + \epsilon$ and if bidder 1 bids any amount greater than this he will win the item.

Is there any kind of one-step procedure that will assign the good to the person with the highest value regardless of the accuracy of the beliefs of the participants? It turns out that the answer is "yes." The *Vickrey auction* works as follows. As before, each person submits a single sealed bid and the item is awarded to the person with the highest bid, but the winning bidder only has to pay the *second-highest* bid. (See [19])

It turns out that the optimal strategy in such an auction is for each person to bid his or her true value for the good. To see this, let b_i be the bid of person i and v_i the true value of person i. For simplicity suppose that there are only two bidders. Then the expected payoff to bidder 1 is

$$\text{Prob}(b_1 > b_2)[v_1 - b_2].$$

If the bracketed term is positive, then bidder 1 wants to make the probability term as large as possible. But if the bracketed term is positive, setting $b_1 = v_1$ makes the probability equal to 1, its maximal value. If the bracketed term is negative, bidder 1 wants to make the probability term as small as possible. But in this circumstance, setting $b_1 = v_1$ makes the probability 0, which is its smallest value. It follows that setting $b_1 = v_1$ is always an optimal strategy. Note that this is a *dominant strategy* in the sense that yields the highest expected payoff to each bidder *regardless* of the other bidder's strategy. Note further that the outcome is essentially the same as the outcome of the standard English auction: the highest bidder gets the item but he pays (essentially) the second highest price.

The Vickrey auction has been used in the computer science literature in [6], [15], [9], and no doubt in several other places. Since it is optimal for each person to reveal his or her true value, the Vickrey auction ensures that the item will be awarded to the person with the highest willingness to pay. However, it *does not* maximize seller revenue: that problem is considerably more complicated since the construction of the revenue-maximizing auction will typically depend on the beliefs of the seller about the buyers' values.

However, it is often the case that the auction that maximizes expected seller revenue has a form similar to that of a Vickrey auction. For example, if there are only two possible valuations for the good, then the seller should set a single take-it-or-leave it price if he believes that there is a high probability that the bidder has the high valuation and otherwise use a Vickrey auction. ([1], page 530.) [11] describes how the New Zealand government used a second-price auction for the spectrum with unfortunate results because they forgot to include this sort of "reserve price" requirement.

2 Computerized agents

The appropriate design of an economic mechanism depends critically on the model that one uses to describe the behavior of the participants. Economists have tended to use game theory to model participant interaction, although there has also been some work with evolutionary models.

Game theory has been justly criticized for its "hyper-rational" view of human behavior. However, such hyper-rationality may actually be an appropriate model for software agents: presumably software agents have much better computational powers than human beings. The whole framework of game theory and mechanism design may well find its most exciting and practical application with computerized agents rather than human agents, a point recognized by [15].

However, there are several additional considerations that come into play with artificial agents rather than human agents. First, to function effectively, a computerized agent has to know a lot about its owner's preferences: e.g., his maximum willingness-to-pay for a good. But if the seller of a good can learn the buyer's willingness-to-pay, he can make the buyer a take-it-or-leave it offer that will extract all of his surplus. Hence *privacy* appears to be a critical problem for "computerized purchasing agents." This consideration usually does not arise with purely human participants, since it is generally thought that they can keep their private values secret.[1]

Secondly, the artificial agent must guard against dynamic strategies that can extract private information. For example, suppose that an agent knows the lowest price at which its master will agree to selling the good (the "reservation price") and that it can safeguard this information from buyers. Suppose that this selling agent follows the simple strategy of accepting any offer that is higher than its reservation price. A buyer can then simply start at 0 and offer a sequence of incremental bids ensure that it ends up purchasing the good at a price slightly more than the seller's reservation price. This will typically not be a good deal for the seller!

This example is far from fanciful. In 1993 the Australian government auctioned off licenses for satellite-television services. The winning bid for one of the licenses, A$212 million, was made by a company called Ucom. Once the government announced Ucom had won, they proceeded to default on their bid, leaving the government to award the license to the second-highest bidder—which was also Ucom! They defaulted on this bid as well; four months later, after several more defaults, they paid A$117 million for the license, which was A$95 million less than their initial winning bid! The license ended up being awarded to the highest bidder at the second highest price—but the poorly designed auction introduced at least a year's delay into pay TV into Australia. See [11] for details of this story and how its lessons were incorporated into the design of the US spectrum auction.

In fact, the example shows why attention to mechanism design is important. If one can construct a mechanism for which truthfully revealing one's true willingness to pay is a dominant strategy, then there is no need to worry about keeping the willingess to pay private. The Vickrey auction is such a mechanism since the dominant strategy in this game is for each party to truthfully reveal the willingness to pay. A mechanism of this sort is called a *direct mechanism*. Somewhat

[1] Even if *current* information can be safeguarded, records of past behavior can be extremely valuable, since historical data can be used to estimate willingness to pay. What should be the appropriate technological and social safeguards to deal with this problem?

surprisingly it turns out that the class of direct mechanisms is much broader than it first appears. A fundamental result in the theory of mechanism design that we will outline below, *the revelation principle*, shows that anything that can be achieved by an arbitrary mechanism can be achieved by a direct mechanism. Hence the issue of keeping the willingness to pay private can be finessed if the mechanism is appropriately designed.

3 A Generalized Vickrey Auction

The Vickrey auction described above is a very powerful mechanism but appears to be of limited scope. However there is a generalization of the Vickrey auction that will handle much more complex problems—including many resource problems that appear to be quit different in nature. The Generalized Vickrey Auction (GVA) that I will describe below appears to be part of the mechanism design folklore, but it doesn't appear to be described in writing anywhere. Here I will provide a detailed argument, but I make no claims of originality except perhaps with respect to the exposition.

Suppose that there are $i = 1, \ldots, n$ consumers who each consume $j = 0, \ldots, k$ goods. Let x_i^j be the consumption of good j by consumer i. Good 0 will denote "money" and $x_i = (x_i^1, \ldots, x_i^k)$ will be the consumption bundle of goods by consumer i. Each consumer i holds some initial consumption bundle $\bar{x}_i$ and some initial amount of money $\bar{x}_i^0$.

An allocation $x = (x_1, \ldots, x_n)$ of goods is *feasible* if the total amount of each good held (including money) equals the total amount available:

$$\sum_{i=1}^n x_i^j = \sum_{i=1}^n \bar{x}_i^j,$$

for each $j = 0, \ldots, k$.

Each consumer i has a utility function $u_i(x) + x_i^0$; this is known as a *quasilinear utility function* and has certain properties that make it convenient for analysis. In particular, there are no "income effects" that influence the demand for the various goods. Note that this allows consumer i's utility to depend on the total allocation of the goods across all consumers not just on how much he gets of each good. In most of our examples, we specialize to the form where $u_i(x) = u_i(x_i)$, but in the last example we use the more general specification.

A reasonable objective in allocating the goods among the consumers is to allocate them in a way that maximizes the sum of utilities:

$$\max_{(x_i^j)} \quad \sum_{i=1}^n u_i(x) + x_i^0$$
$$\sum_{i=1}^n x_i^j = \sum_{i=1}^n \bar{x}_i^j$$
$$\text{for all } j = 0, \ldots n$$

In the simple case of the Vickrey auction described above, the utility functions were simply the difference between the value, v_i and the payment made by the consumer. Just as the consumer will not want to reveal his value to the seller, the participants in this resource allocation problem will not in general want to reveal their true utility functions. Our problem is to design a mechanism that will induce the participants to truthfully reveal their private information.

The Generalized Vickrey Auction

1. Each consumer i reports a utility function $r_i(\cdot)$ (which may or may not be the truth) to the center.

2. The center calculates the allocation (x_i^*) that maximizes the sum of the reported utilities subject to the resource constraint.

3. The center also calculates the allocation $(\hat{x}_{\sim i})$ that maximizes the sum of the utilities *other* than that of consumer i subject to the constraint that the allocation not use any of consumer i's resources.

4. Agent i receives the bundle x_i^* and receives a payment of $\sum_{j \neq i}[r_j(x^*) - r_j(\hat{x}_{\sim i})]$ from the center.

The final payoff to agent i in the GVA is given by

$$u_i(x^*) + \sum_{j \neq i} r_j(x^*) - \sum_{j \neq i} r_j(\hat{x}_{\sim i}).$$

I claim that if the GVA mechanism is used then it is in the interest of consumer i to report his true utility function $r_i(\cdot) = u_i(\cdot)$.

The first step in the argument is to note that the third term in the sum is irrelevant to consumer i's decision since it is totally outside of his control. To emphasize this we denote this term by K. It is useful in reducing the magnitude of the sidepayment to consumer i, but has no effect on the strategy of consumer i.

Next observe that the center will choose x_i^* to maximize

$$r_i(x) + \sum_{j \neq i} r_j(x)$$

subject to the resource constraint and consumer i wants them to maximize his payoff,

$$u_i(x) + \sum_{j \neq i} r_j(x) - K.$$

By inspection of these two equations it is optimal for the consumer to report $w_i(\cdot) = u_i(\cdot)$. This concludes the argument.

4 Examples of the GVA

Here we examine a few special cases of the GVA.

The Standard Vickrey Auction. In this case, the utility function of consumer i is $v_i - p$, where v_i is consumer i's value and p is the price he pays. Let $x_i = 1$ if consumer i gets the good and $x_i = 0$ if he does not. Then the sum of the utilities becomes

$$\sum_{i=1}^{n} v_i x_i$$

and the resource constraint is

$$\sum_{i=1}^{n} x_i = 1.$$

Of course x_i must be an integer, but this ends up being satisfied automatically so there is no need to impose that as an additional constraint.

Let m be index the consumer with the maximum value of v_i; then in order to maximize the sum of utilities the center will allocate $x_m^* = 1$ and $x_j = 0$ for all $j \neq m$. Let consumer s have the second-highest value; then if we eliminate consumer m the maximal sum of the remaining utilities will be v_s. The net payoff to consumer m in the GVA will be $v_m - v_s$ which is exactly the same as the Vickrey auction.

Multiple units of the good. Suppose that there is one good but there are $\bar{x}$ units of it to sell. Let (x_i^*) be the allocation that maximizes the sum of all consumers' utilities and let $\hat{x}_{j \sim i}$ be the amount allocated to consumer j if the sum of all consumers' utilities but consumer i

is maximized. Then consumer i's payoff is

$$u_i(x_i^*) + \sum_{j \neq i} u_j(x_j^*) - \sum_{j \neq i} u_j(x_{j\sim i}).$$

To see how this works, suppose that there are 2 consumers and 3 units of the good to allocate. Consumer 1 values the first unit of the good at 10, the second unit at 8 and the third unit at 5. Consumer 2 values the goods at (9, 7, 6), respectively. By inspection the optimal assignment is to give consumer 1 two units of the good and consumer 2 one unit of the good. Consumer 1 receives a total utility of 18 and consumer 2 receives a total utility of 9.

Here's how the GVA handles this problem. If consumer 1 isn't present, all the goods go to consumer 2 who receives a utility of $9 + 7 + 6 = 22$. In the GVA, Consumer 1's net payoff is

$$18 + [9 - 22] = 18 - 13 = 5.$$

So consumer 1 pays 13 for the 2 units of the good he receives.

Similarly, if consumer 2 isn't present, all the goods go to consumer 1 who receives a utility of $10 + 8 + 5 = 23$. Consumer 2's net payoff is then

$$9 + [18 - 23] = 9 - 5 = 4.$$

Hence consumer 2 pays 5 for the 1 unit of the good that he receives. The seller receives $13 + 5 = 18$ for the 3 units that he has sold.

Public goods. Suppose that each consumer i initially owns $\bar{x}_i$ units of the good. Consumer i can contribute x_i to a "collective good" (e.g., a pool for site licensed software) which will result in a total collection of $G = \sum_{i=1}^n x_i$. The sum of utilities is

$$\sum_{i=1}^n u_i(G) + \sum_{i=1}^n \bar{x}_i - G.$$

We assume that $u_i(\cdot)$ is a differentiable, increasing, concave function.

This is a classic public goods problem. The G^* that maximizes the sum of utilities satisfies the condition

$$\sum_{i=1}^n u_i'(G^*) = 1,$$

whereas the contribution that is optimal for each agent i acting on his own satisfies the condition

$$u_i'(G^\dagger) = 1.$$

Under the conditions we have assumed, the total voluntary contributions will be smaller than the socially optimal amount.

How does the GVA work to solve this problem? Let (x_i^*) be a pattern of contributions that maximizes the sum of utilities and let $(\hat{x}_{j\sim i})$ be the pattern of contributions that maximizes the sum of utilities omitting the utility and contribution of consumer i. The payoff to consumer i is then

$$u_i(x_i^*) + \sum_{j \neq i} [u_j(x_j^*) - u_j(\hat{x}_{j\sim i}].$$

To see how this works in practice suppose that there are three consumers each with an initial wealth of 10. If the total contributed to the collective good is $G = x_1 + x_2 + x_3$, consumer i will have a net value of $.4G - x_i$. The sum of the utilities over all 3 consumers is

$$1.2G + (30 - G) = 30 + .2G,$$

which is clearly maximized when $x_1 = x_2 = x_3 = 10$. The sum of utilities over any 2 consumers is

$$.8G + (20 - G) = 20 - .2G,$$

which is maximized when $x_1 = x_2 = x_3 = 0$. Hence the equilibrium payoff to consumer i is

$$[.4\times 30-10]+[.8\times 30-20]-[.8\times 0-20] = 26.$$

Effectively each consumer must pay $.8 \times 30 = 24 - 20 = 4$ as a "tax" on top of the payment of 10 that he is already making. This tax represents the cost that the consumer is imposing on the other consumers through his presence. To see this note that if consumer 1 isn't present, the public good will not be provided. Therefore, it is the presence of consumer 1 that imposes a "cost" of $.2 \times 10$ on each of the other consumers.

This mechanism is essentially the celebrated Groves-Clarke mechanism ([5], [2]). In fact the proof of the GVA presented earlier is essentially the standard proof of Groves-Clarke result. (See, e.g., [18], p. 429.) The interesting fact is that this standard argument works for a much broader class of resource allocation problems than the classic public goods problem to which it is normally applied.

5 The Revelation Principle

The GVA is called a *direct revelation mechanism* since the "message" sent to the center is in fact the entire private information of the consumer: his utility function. One might imagine other "indirect" mechanism: the consumer announces a bid, or a reservation price. It is rather remarkable that anything that can be achieved by such an "indirect" mechanism can be achieved by a direct mechanism. This assertion is known as the *revelation principle*.

In this paper we have considered only mechanisms for which truth-telling is a *dominant strategy*. The revelation principle is valid under much more general circumstances but it is particularly easy to explain in this case.[2]

For notational simplicity let us index the different types of utility function by t so that $u_i(t, x)$ is consumer i's true utility if the outcome is x. Let r_i be agent i's reported utility type, let $r = (r_1, \ldots, r_n)$ be the set of all reports, and let $x(r)$ be the outcome if the reports are r. The function that assigns the outcome $x(r)$ is the mechanism. If truthtelling is a dominant strategy for each agent i then it must be the case that

$$\begin{gathered} u_i(t, x(r_1, \ldots, t_i, \ldots, r_n)) \\ \geq u_i(t, x(r_1, \ldots, r_i, \ldots r_n)) \\ \text{for all reports } r_i \end{gathered}$$

(This is called the *incentive compatibility constraint.*)

Let us now consider some other mechanism. Rather than just reporting the type t, this other mechanism allows consumer i to send some different message, m_i. If the consumers send messages $m = (m_1, \ldots, m_n)$ the resulting allocation is y. If m_i^* is a dominant strategy for consumer i

$$\begin{gathered} u_i(t, y(m_1, \ldots m_i^*, \ldots, m_n)) \\ \geq u_i(t, y(m_1, \ldots m_i, \ldots, m_n) \\ \text{for all messages } m_i \end{gathered}$$

What can consumer i's message depend on? It can't depend on the the other consumers' types since consumer i doesn't

[2] In fact, the revelation principle was first formulated for dominant strategy equilibria by [4].

know them. All that consumer i's message can depend on is his private information—i.e., his type. Accordingly, let us define a function $M_i(t) = m_i^*$ that gives the optimal message for consumer i if his type is t. By definition, $M_i(t)$ must satisfy

$$u_i(t, y(m_1, \ldots, M_i(t), \ldots, m_n)) \geq$$
$$u_i(t, y(m_1, \ldots m_i, \ldots, m_n)$$
$$\text{for all messages } m_i,$$

which is exactly the condition that characterizes a direct revelation mechanism. Since the optimal message only depends on the *true* type there is no loss in generality in designing the mechanism so that the message *is* the type.

In other words there is no loss of generality in restricting ourselves to direct revelation mechanism. This is very important for the design of computerized agents since it says in effect that there is nothing to be gained (or lost) by communicating anything other than the "essentials" of the problem.

Consider, for example, the auction problems examined earlier. Each consumer had an incentive to reveal his true value v_i; the auction design itself ensured that the consumer was not hurt by this full revelation. The fact that we can restrict ourselves to direct mechanisms makes the privacy issue alluded to before much less troublesome.

6 Brief introduction to the literature

The classic work that laid out the rationale and basic framework for the field of mechanism design is [7] . Useful surveys of mechanism design are available in [3]; [14] and [8] are particularly useful.

For interesting applications of mechanism design see [12], [13], [10], [11], and [20] for auction design. See [16] for matching models and [17] for price discrimination.

Several computer science applications of mechanism design that were influenced by the mechanism design literature are described in [15].

References

[1] Ken Binmore. *Fun and Games*. D. C. Heath, Lexington, MA, 1992.

[2] E. Clarke. Multipart pricing of public goods. *Public Choice*, 11:17–23, 1971.

[3] John Eatwell, Murray Milgate, and Peter Newman. *Allocation, Information and Markets*. W. W. Norton & Co., New York, 1989.

[4] Alan Gibbard. Manipulation of voting schemes: a general result. *Econometrica*, 41:587–602, 1973.

[5] Theodore Groves. Incentives in teams. *Econometrica*, 41:617–631, 1973.

[6] Bernardo Huberman and Tad Hogg. Distributed computation as an economic system. *Journal of Economic Perspectives*, 9(1):141–152, 1995.

[7] Leo Hurwicz. The design of mechanisms for resource allocation. *American Economic Review Papers and Proceedings*, 63:1–30, 1973.

[8] John O. Ledyard. Incentive compatibility. In John Eatwell, Murray Milgate, and Peter Newman, editors, *Allocation, Information and Markets*.

W. W. Norton & Co., New York, 1989.

[9] Jeffrey K. MacKie-Mason and Hal R. Varian. Pricing the Internet. In Brian Kahin and James Keller, editors, *Public Access to the Internet*. MIT Press, Cambridge, MA, 1995.

[10] Preston R. McAfee and John McMillan. Auctions and bidding. *Journal of Economic Literature*, 25:699–738, 1987.

[11] John McMillan. Selling spectrum rights. *Journal of Economic Perspectives*, 8(3):145–162, 1994.

[12] Paul Milgrom. Auction theory. In Truman Bewley, editor, *Advances in Economic Theory, 1985: Fifth World Congress*, pages 1–32. Cambridge University Press, 1985.

[13] Paul Milgrom. Auctions and bidding: a primer. *Journal of Economic Perspectives*, 3(3):3–22, 1989.

[14] Roger B. Myerson. Mechanism design. In John Eatwell, Murray Milgate, and Peter Newman, editors, *Allocation, Information and Markets*. W. W. Norton & Co., New York, 1989.

[15] Jeffrey S. Rosenschein and Gilad Zlotkin. *Rules of Encounter*. MIT Press, Cambridge, MA, 1994.

[16] Alvin E. Roth and Marilda A. Oliviera Sotomayor. *Two-Sided Matching*. Cambridge University Press, Cambridge, England, 1990.

[17] Hal R. Varian. Price discrimination. In Richard Schmalensee and Robert Willig, editors, *Handbook of Industrial Organization*. North-Holland Press, Amsterdam, 1989.

[18] Hal R. Varian. *Microeconomic Analysis*. W. W. Norton & Co., New York, 1992.

[19] William Vickrey. Counterspeculation, auctions and competitive sealed tenders. *Journal of Finance*, 16:8–37, 1961.

[20] Robert Wilson. Auction theory. In J. Eatwell, M. Milgate, and P. Newman, editors, *The New Palgrave*. MacMillan, London, 1987.