Economic Perspectives—Volume 1, Number 2—Fall 1987—Pages 113–129

Information and the Coase Theorem

Joseph Farrell

Every economist kept awake by noisy neighbors must have relieved the tedium of counting sheep by pondering the social institutions that make this kind of thing happen. A classical answer is that Pareto-efficiency normally requires competitive markets, and since there is obviously no such market for "quiet in my house tonight," inefficiency should be no surprise. In 1960, Ronald Coase's seminal paper "The Problem of Social Cost" challenged that view. Its claim is that complete competitive markets are not necessary for efficiency. Rather, if the market outcome is inefficient, then people will get together and negotiate their way to efficiency.

Although Coase was concerned with how negotiation can repair externalities, the argument is much broader. It says that if nothing obstructs efficient bargaining then people will negotiate until they reach Pareto-efficiency. This claim is far more ambitious than the traditional competitive-equilibrium welfare theorems. For instance, as Calabresi (1968) pointed out, it implies that we should not worry about monopoly or the provision of public goods either: people can negotiate their way to efficiency.

Of course, it is a tautology that *if* people negotiate efficiently *then* every outcome will be efficient (else people would negotiate something better). The Coase theorem is important only if we believe that efficient negotiation is likely. Some economists think the Coase theorem implies a lot about the proper scope of government intervention in the economy and about the welfare consequences of laissez-faire. Others see it as a

■ *Joseph Farrell is Visiting Assistant Professor of Economics, University of California–Berkeley, Berkeley, California.*

mere tautology: of course we can attribute inefficiencies to "bargaining imperfection," but it may not be useful to do so. Such attitudes sometimes seem to depend more on ideology than on reason. This article gives one economist's reasoned view.

The Coase Theorem as a Decentralization Result

Statements like, "If such-and-such conditions hold, then selfishly optimal individual decisions will lead to efficient aggregate outcomes," are called *decentralization results*. For instance, the familiar "welfare theorem" says that any competitive equilibrium with complete markets is Pareto-efficient. But no real economy has the required competitive markets in every good (indexed by time, place, and state-of-the-world). Some of the structural causes of this (like economies of scale) are straightforward, and so economists understand fairly well when the welfare theorem applies and when it does not.

On first acquaintance, the Coase theorem seems much more robust. Like the welfare theorem, it says that if everything is tradeable then Pareto-efficient outcomes result. Unlike the welfare theorem, it makes no strong assumptions about convexity, price-taking, and complete markets. Instead, a one-line argument says that, absent barriers to contracting, all must be well!

But as the U.S. Postal Service warns its customers, "If it seems too good to be true, it probably is." The Coase theorem dispenses with the heavy assumptions of perfect competition, but replaces them with the strong assumption that no mutually beneficial agreement is missed. So while it economizes on formal institutions, it demands a lot of coordination and negotiation. We can see this sharply by rephrasing the welfare theorem: in competitive markets, if a market-clearing outcome is Pareto-inefficient, then some agent (consumer or firm) is not at a maximum. That is, *unilateral* action can improve on anything inefficient. Without competitive markets, even when each individual is maximizing, the outcome is often inefficient. Improvements demand joint action, which requires coordination and negotiation. In this sense, competitive markets solve our coordination problems.

Looking at things this way suggests that we try to assess the difficulty of coordination and negotiation between the people who must get together to improve on inefficient outcomes. For instance, if many people must cooperate, it is harder than if only a handful are involved. Getting many people to negotiate successfully about supplying a public good is hard. (And one could argue that government is the right institution to deal with such problems, so that the Coase theorem loses its decentralization bite.) In this article, for simplicity, I will deal only with problems involving pairs of people—the bilateral externality problems of a pair of neighbors, for instance.

How, then, can we evaluate the claim that in bilateral negotiations rational economic people are likely to emerge with relatively efficient outcomes? Most expositions of Coase's argument focus on the inefficiencies that arise if tradeable property rights are not clearly established, and then suggest that any other inefficiencies must result from imperfections in bargaining, due perhaps to holdouts or to transactions costs. But these vague terms are not very helpful, and, in practice, bargaining problems may be as important a bar to efficiency as missing property rights. We ask, then, what are the causes of imperfections in bargaining, how policy affects transaction costs, and when these problems are severe compared with some alternative such as central direction.

Recently, many economists have analyzed the efficiency of bargaining. The so-called "cooperative" or "axiomatic" theory of bargaining is not helpful here, since one of its axioms is that outcomes are efficient! But a large literature has developed on "non-cooperative" models of bargaining, formalizing the process of offers and counteroffers that we see in real negotiations. The main conclusion of these models (recently surveyed by Sutton, 1986) is that bargaining is typically inefficient when, as is likely, each bargainer knows something relevant that the other does not, such as his payoff from a successful agreement. The inefficiencies consist of bargains not struck that should be, excessive delay and other direct costs of bargaining. Typically, each bargainer incurs and imposes real costs to change the likely price to his advantage.

If everyone knew all about everyone else, it is hard to envision how negotiation could drag on or break down. This intuition that bargaining will be efficient when everyone's tastes and opportunities are common knowledge is borne out by some recent studies. For instance, Rubinstein (1982) shows that, in one natural bargaining game with symmetric information, bargaining ends at once with one agent making an offer that the other is (just) willing to accept. Roth and Murnighan (1982) report experiments confirming that when bargainers' payoffs are common knowledge, little disagreement results. But when people don't know one another's tastes or opportunities, then experience, theory and experimental evidence all confirm that negotiations may be protracted, costly and unsuccessful. A potential buyer may value a house more than its prospective seller does, but less than the seller believes "most" buyers do. He would then have trouble persuading the seller to lower the price enough to make the deal. Exactly the same is true if it is quiet, rather than a house, that is being bought and sold.

So we cannot assume that all mutually beneficial contracts are signed, unless we assume that everyone knows everything about everyone, which they do not. The strong form of the Coase theorem—the claim that voluntary negotiation will lead to fully efficient outcomes—is implausible unless people know one another exceptionally well. As I shall argue, that case is not only unlikely, but is also the case where decentralization is least useful.

Why Are Decentralization Results Interesting?

Decentralization results are of interest for at least three reasons. First, as Adam Smith emphasized, they are surprising. Analyzing (in game theory) the aggregate results of individuals' selfishly-optimal choices, we find that those results are typically not Pareto-efficient. It is remarkable that a model with any resemblance to our economy predicts efficiency. Second, decentralization results give us a taxonomy of inefficiency: for instance, the welfare theorem lets us classify inefficiencies as due to monopoly, externalities, and so on. This helps us to understand and perhaps to solve such inefficiencies, just as a doctor's diagnosis (what is different about this patient from a healthy patient?) is part of treatment. Third and perhaps most important, people often use decentralization results—expecially the Coase theorem—as arguments against government intervention. They might claim, for instance, that the neighbors should be expected to reach an efficient outcome through private negotiations, without outside intervention.

As arguments against active government policy, the welfare theorem and the Coase theorem are unconvincing. They say that, in ideal circumstances, the laissez-faire outcome is no less Pareto-efficient than the ideal government-dictated outcome. But they do not claim that it is better; further, centralization has some obvious advantages, as in problems of equity. Why, then, do so many economists see the welfare theorem and the Coase theorem as powerful arguments against intervention? Presumably they think that the market process in practice comes closer to ideal performance than does the actual process of government. But this belief can not be proven by analyzing models of the market and of government intervention that both give efficient outcomes. Instead, the imperfections of each system must be modelled.

A common complaint about centralized decisions is that they cannot properly adjust to the special circumstances of each case, as decentralized decisions can. Formally, people have private information that should affect decisions, and that for some reason is not available to a central authority. Hayek (1945) was perhaps the first to emphasize such informational problems. He argued that the problem solved by the welfare theorem is "emphatically *not* the economic problem which society faces." Society's problem is to make the best use of its knowledge, which "never exists in concentrated or integrated form, but solely as the dispersed bits of incomplete... knowledge which all the separate individuals possess." The challenge of decentralization is "whether we are more likely to succeed in putting at the disposal of a single central authority all the knowledge which ought to be used... or in conveying to the individuals such additional knowledge as they need." From Hayek's point of view, the Walrasian "decentralization" result is no such thing, since the central auctioneer must collect all (or a vast amount of) the economy's information to set market-clearing prices, and such an economy is thoroughly centralized. True decentralization consists in delegating decisions to those who know more about them. If there were no private information, to take an extreme version of Hayek's view, decentralization would have no point: a central authority would be perfectly able to take fully efficient decisions. Coase points out that decentralized negotiation would also work well then, but that is not a very exciting observation. It is much more instructive to compare how different systems do when private information does exist. Modern analysis of bargaining under incomplete information shows that property rights and negotiation will not lead to fully efficient outcomes in that interesting case (Samuelson, 1985). So the Coase theorem's strong claim, viewed as a decentralization result, is false when it is most interesting. But before dismissing the argument, let us be clearer about what is being compared.

Mechanism Design: Getting Private Information Revealed

Property rights and negotiation will not yield first-best outcomes when there is important private information, and that case is the one that should be examined. So property rights do not give fully efficient decentralization in real or interesting problems. But perhaps first-best efficiency is too demanding a standard; perhaps no social arrangement is that good when private information is important and people's goals differ. If so, then it would be silly to berate property rights for their failure.

Private information poses more than one problem for efficient choice. We shall return below to Hayek's view that no central authority can cope with the complexity of all the relevant information. Before doing so, we discuss another problem of central decision making, one which Hayek ignored: one must give people incentives to reveal what they know, assuming (contrary to Hayek) that the central authority can cope if they do so. The study of such incentives is the theory of *mechanism design*. Mechanism design captures some important virtues of central authority but misses its flaws.

People with private information may not readily reveal it, especially if they know that it will be used in a decision that affects them. Unless everyone shares the same goals, people typically have incentives to lie. For instance, suppose the government asks people how much they would value a public project, and plans to do the project if people seem keen enough. Then anyone who values the project more than he expects to be taxed for it will be tempted to exaggerate his enthusiasm, so as to maximize the chance that the government will go ahead; others (who would value it, but not enough) will pretend that they don't want it at all. Is there some way to make people's tax payments depend on their declared values and thus to cure those incentive problems? In other words, can one persuade people to tell the truth when they know how the information will be used? That is the subject of mechanism design.[1]

The formal framework is this. There is a central authority, whom I will call the king. He must make some decision, and to make a good decision he needs some facts that other people (his "subjects") know. Because the subjects care about the decision, but their goals differ from his, he must give them incentives to tell the truth. To do so, the king can commit himself to an incentive scheme: formally, this scheme specifies how the decision, and perhaps some money payments, will depend on the reported information.

Sometimes, to elicit important information, the king may have to promise to do things that he may not want to do once the information comes out, and so the

[1]For an introduction, see, for instance, Sonnenschein (1983), or for a more complete survey, Laffont and Maskin (1982).

assumption that he can commit himself is important. For instance, in an early incentive problem, Solomon had to decide which of two women was in fact the mother of a living infant boy whom they both claimed. (The other's son had died.) His solution is reported in First Book of Kings, Chapter 3, as follows:

> Then the king said, "The one says, 'This is my son that is alive, and your son is dead'; and the other says, 'No; but your son is dead, and my son is the living one.'" And the king said, "Bring me a sword." So a sword was brought before the king. And the king said, "Divide the living child in two, and give half to the one, and half to the other." Then the woman whose son was alive said to the king, because her heart yearned for her son, "Oh, my lord, give her the living child, and by no means slay it." But the other said, "It shall be neither mine nor yours; divide it." Then the king answered and said, "Give the living child to the first woman, and by no means slay it; she is its mother." And all Israel heard of the judgment which the king had rendered; and they stood in awe of the king, because they perceived that the wisdom of God was in him, to render justice.

Solomon's solution worked because the impostor apparently would have preferred a dead child to ceding it, or perhaps because she failed to see what the king was doing. Otherwise she would have done better to say what the real mother said, and take a chance of getting the boy. So perhaps Solomon was lucky. Or could he rely on finding some clever scheme?

Mechanism design theory answers this question for us, and broadly the answer is yes. A wise king like Solomon can find schemes to get the first-best outcome in trickier problems than the disputed baby. In general, doing so requires some *side payments*, which help establish people's true willingness to pay for particular outcomes, and thus show what decision would maximize net benefits. For instance, if the boy's mother values him more than does the impostor, then Solomon could have got him to the right mother by confiscating him and then auctioning him off! In much more general problems, Groves and Ledyard (1977), d'Aspremont and Gerard-Varet (1979), and others have shown how Solomon can set up a cunning scheme that makes side payments depend on reported information, so as to achieve fully efficient outcomes without losing the incentives to tell the truth.

The idea is that when you claim to want something very much, that claim will very likely make Solomon do what you ask. For instance, if you claim that tonight of all nights is an especially good time to turn up the Talking Heads, Solomon is likely to say "OK." But since that may be the opposite of what other people want, you should pay for the effect on them of this increased chance of the decision going your way; you should pay the *expected externality* that you create by your claim of intense preference. This is not for reasons of justice or equity (though those may apply too); it is because only if you must make such a payment will you have the right incentives to claim intense preference when, and only when, you feel it. By paying for the effects of your claim on others' expected welfare, you internalize the whole social problem when you report your private information, so it's not surprising that you have all the right incentives then. In this example, if you must pay for your neighbor's lost sleep, you will only tell Solomon you *must* have a party when in fact your urge is especially intense; similarly, your neighbor will only claim that he *must* get a good night's sleep when in fact he really needs it, since he must pay for your lost party.

Another way of looking at this problem is that each person can (by changing his report) get the others to give in to him in some way, by paying an amount just equal to their true reservation price for giving in; that is what the expected externality measures. So the scheme amounts to an idealized process of bargaining, stripped of strategic holdout problems: that is, a process in which everyone makes concessions at cost. Of course, such a process leads to efficient outcomes. But while Coase suggested that the efficiency of ideal bargaining means that everything can be decentralized, the mechanism-design view is that it means the opposite: centralization lets us have such a process (through an expected-externality scheme) while we know that decentralized bargaining is imperfect when there is private information.

This analysis of centralized authority, then, is optimistic about what a wise and benevolent central planner can do with suitably cunning schemes. What can we say for decentralization? There are two kinds of decentralization, each of which has important advantages in reality, but in this model they only impair efficiency. First, under *administrative decentralization*, instead of sending all information to the center (the king), we delegate some decisions to "princes" who have only partial information. In the mechanism-design model, such delegation cannot possibly be useful; anything the princes can do, the king can at least as well. (If the king gets all the information that the princes would get, then he can work out what they would do, and do the same. And if people are not tempted to lie to the princes, then they will not lie to the king if he promises to do that.) Thus decentralization becomes a mere subset of centralization, and is at best harmless. For instance, Solomon might be willing to delegate to the music-lover the choice of Mahler versus Bruckner. But if it were Scarlatti versus the Stones, it might affect how late Solomon would want to let you play music, so delegation might be troublesome. In general, since the social decision problem can't be split up into chunks with no interdependence, such administrative decentralization is not harmless. Of course, this observation points up the idealized view that mechanism-design takes of the central administrator; we will return to this point below.

A second, more radical, concept is *political decentralization*, in which people are always free to choose whether to participate in any relationship with others. Voluntary trade has a good press among economists (see for instance Friedman and Friedman, 1979), and it may be surprising that it often mars efficiency. The reason is a little more subtle than the obvious drawback to administrative decentralization that we just described, and it is closely related to the timing of commitments.

If people make voluntary agreements when they know nothing that is not common knowledge, then the freedom to leave just ensures that everyone shares in the benefits from cooperation: voluntarism has only distributional effects. Distributing the ex-ante gains somehow, people can sign contracts that bind them to essentially the same schemes that Solomon would impose on them. There is really no difference between such ex-ante contracts and central authority; indeed, voluntary contracts can include the appointment of a mediator to collect and process information, and to say what should be done.

But if people sign voluntary contracts *already knowing something that others do not know*, then things are different. In that case, voluntary contracts must not only be designed to prevent cheating, but also must give every "type"[2] of person positive surplus, or else he won't participate; and in general it is costly to efficiency if some types refuse to play. For instance, the confiscate-and-auction solution to the problem of who-should-get-it runs into trouble here; the owner either loses his object or else has to buy what he already had, so he'd rather just walk off. Similarly, Solomon's scheme to allocate the baby would not have worked if the impostor could not have been forced to the king's court.

A contract that makes every player willing to participate *once he knows his own private information* is called "individually rational." By definition, voluntary schemes must respect this constraint, but schemes imposed by the king need not. Since it imposes an extra constraint on the available schemes, this political decentralization obviously could damage efficiency. But it is perhaps surprising just how widespread that damage is. For instance, Myerson and Satterthwaite (1983) have shown that, in a class of bargaining problems, the first-best outcome is not attainable in any individually rational contract. That is, if people come to bargaining already knowing their private values for a good, then no arrangement exists that will lead them to trade precisely when they should, given that each can choose to walk away. So the king's power to coerce really helps to achieve efficiency.

For instance, suppose the problem is which of two people should have an indivisible object, a "seller" (who originally has it) or a "buyer." The efficient solution is that whoever in fact values it more should have it, with perhaps some payment to the other. (That is, every Pareto-efficient outcome has this form.) The king can easily achieve this outcome using an incentive-compatible scheme if participation is compulsory: for example, he can confiscate the item from the "seller" and then auction it off, dividing the revenues equally between the two people. But this solution is not feasible with voluntary trade; the seller may prefer to keep the object rather than to participate and risk having to repurchase (or lose) something he already has. A lump sum payment to the seller could solve that problem, but then the buyer (who would have to make that payment) might prefer to withdraw. And payments to encourage participation conditional on reported "type" (value) would upset the incentive properties of the confiscation/auction scheme. Myerson and Satterthwaite, in fact, show that there is *no* individually-rational incentive-compatible scheme that always

[2]A "type" of person is the possible participant who knows that his private information takes a particular value. For example, in a bargaining problem where the seller's reservation price is known to him alone, one type of seller is the seller who knows that his reservation price is $5.

yields efficient outcomes in such a problem if it is unknown who values the object more. The same argument applies to our difficult neighbors: since we cannot say a priori what time people should go to sleep, voluntary bargaining cannot reproduce the good results that Solomon could get with a compulsory-participation scheme.

Thus the fact that voluntary exchange must make every type of participant better off actually hurts its ex-ante efficiency. Allowing some chance of harming one party can make each party substantially better-off in expected value. Political decentralization finds no favor in this mechanism-design outlook.

Does Mechanism Design Miss the Point? Back to Hayek

These arguments do tell us something important about the virtues of centralization. Central authority helps when decisions are so interdependent that they cannot well be delegated; and it can also help efficiency by making recalcitrant people participate in schemes that benefit society in general. But it is not really plausible that central authority always dominates decentralization; the arguments above do not capture the problems of centralization.

What are those problems? I will simply mention one and discuss another in more detail. One problem arises if people do not trust the king's commitment to an incentive scheme. For instance, some people think that the IRS should not be allowed to share computer files with other government agencies, lest someone be unable to resist the temptation to use information in an inappropriate way. And if a central authority fails to succeed in committing himself to ignore information, it can lead to excessive incentives for subjects to try to influence his decisions; this is the idea of "rent-seeking" behavior, which often has important social costs.

Another problem, which Hayek emphasized, is whether the king can handle the job of collecting and using the relevant information regardless of the incentives. People are not very good at processing information, and computers are notoriously untrustworthy outside routine tasks. Hayek argued that much important economic knowledge is "of the kind which by its nature cannot enter into statistics and therefore cannot be conveyed to any central authority..." He added: "We cannot expect that [our] problem will be solved by first communicating all this knowledge to a central board which, after integrating *all* knowledge, issues its orders... We need decentralization because only thus can we ensure that the knowledge of the particular circumstances... will be promptly used." And later, "the problem is precisely how to extend the span of our utilization of resources beyond the span of the control of any one mind."

All this, if we believe it, suggests that to evaluate decentralization under incomplete information we should analyze a model in which the king has a limited ability to process information. In other words, we need to recognize that centralized schemes must be relatively simple, in the sense of ignoring much relevant information.

A Second-Best Coase Theorem? Property Rights Versus the Bumbling Bureaucrat

Property rights and voluntary private negotiation fail to achieve "first-best" efficient outcomes when there is important private information. And such outcomes often can be achieved, despite the information problems, by a wise and benevolent king who is prepared to coerce people to participate in an incentive scheme. But there has been no such king since Solomon. Central authority may be better portrayed by the image of a bureaucrat, perhaps benevolent but certainly bumbling. Might not property rights and private negotiation look better next to such a bumbler than next to Solomon? In his 1960 article, Coase argued for some kind of second-best comparison like this, as against comparing things with first-best efficiency, so perhaps the Coase theorem should be viewed as a second-best result: property-rights are more efficient than some reasonable alternative. In this section I use a simple mathematical example to make such a comparison. Thinking of our sleepy and noisy neighbors, we want to compare the likely efficiency of their imperfect negotiation with that of a city ordinance that bans noise after, say, 10 p.m., and thereby implicitly allows it up to that time.

Suppose that a decision x (a real number: 10 p.m. or 9:30 p.m. or when?) must be taken, and that two people, A and B, care about it. Each privately prefers some value for x: A would like $x = a$, and B would like $x = b$, where $a < b$. Each dislikes deviations of x from his preferred value. In fact, we represent payoffs by

$$u(x, a) = -\alpha(x - a)^2 \qquad (A\text{'s payoff})$$

$$v(x, b) = -\beta(x - b)^2 \qquad (B\text{'s payoff}).$$

We suppose that these payoffs are represented in dollar terms, so that any side payments can simply be added to or subtracted from u and v; we also suppose that A and B are risk-neutral. The utility functions u and v are common knowledge, as are the parameters α and β, which represent the importance of the choice to A and to B, and which for convenience we assume sum to 1. However, only A knows a and only B knows b. To A and to the outside world, b is uniformly distributed on an interval $[b_-, b_+]$; and similarly a is (independently) uniform on $[a_-, a_+]$; we assume that $a_+ < b_-$. It is convenient to define $E(a)$ as the expected value of a, $E(a) = [a_- + a_+]/2$, and likewise $E(b) = [b_- + b_+]/2$. We write C for the expected degree of conflict, $E(b) - E(a)$. Finally, we write r for the variance of a, $r = [a_+ - a_-]^2/12$, and s for the variance of b, $[b_+ - b_-]^2/12$.

If a and b were public, then Pareto–efficiency would simply mean choosing x to maximize $u + v$ (recall that u and v are expressed in dollar terms), or in other words to minimize $\alpha(x - a)^2 + \beta(x - b)^2$. A little algebra shows that this expression is equal to $(x - x^*)^2$ plus a constant independent of x, where $x^* = \alpha a + \beta b$. So efficiency would simply require that $x = x^*$. We call this the "first-best" solution to our choice-of-x problem. Notice that x^* depends on a and b, so that the private information really is relevant.

King Solomon would have little trouble finding x^*, even though a and b are private information. Guided by the modern theory of mechanism design, he would ask A and B to tell him a and b, promising first that he would order x to be set equal to x^*, calculated on the assumption that A and B tell the truth; and second that A and B would have to pay each other, or the treasury, sums of money that depend on their reported values of a and b respectively. If A reports that $a = a'$, for instance, then he has to pay the expected value (where b is the unknown) of $\beta(\alpha a' + \beta b - b)^2$, which is the net effect on B's payoff of A's reporting a'.[3] Similarly, B must pay the expected net effect of his report b' on A's welfare. As a result, each person internalizes the whole social payoff, and so each has incentives (as the reader can readily check) to report accurately: to set $a' = a$ and $b' = b$ respectively.

So Solomon gets x^*. But the bumbling bureaucrat is not up to Solomon's standard, and cannot handle such a scheme. He must make his decision based only on public information. Because u and v are quadratic, his best choice is to set x at $x^B = \alpha E(a) + \beta E(b)$. This achieves a good compromise between A's and B's interests: if a and b happen to be at their expected values $E(a)$ and $E(b)$ then it is fully optimal. But because the bureaucrat can use only public information, his decision cannot respond to variations in a and b around their means. The resulting loss in welfare, evaluated from the point of view of someone (like ourselves) who knows neither a nor b, is $\alpha^2 r + \beta^2 s$: it is variance in a and b that makes the bureaucrat inefficient. So we can assess his imperfection (compared to Solomon) at $\alpha^2 r + \beta^2 s$.

We now evaluate the "property rights" system that gives one of the parties (say, A) the right to choose x, but lets B offer bribes to affect A's choice of x. In other words, the two parties can sign a contract that specifies x and gives A money from B. We impose no restrictions on the complexity or enforceability of that contract, and no transaction costs on the parties in negotiating it. In fact, we suppose (for definiteness) that B offers a "menu" of bribes in return for different possible choices of x. What happens?

There *is* a contract that would always get A to choose the most efficient value of x: it specifies that B will pay A $p(x) = \beta[(a_- - b)^2 - (x - b)^2]$ if A chooses x. Thus A internalizes all the social gain from a responsible choice of x. But while this contract (essentially uniquely) maximizes joint surplus, B actually does worse than with no contract (he emerges with $-\beta(a_- - b)^2$, which is worse than the expected value of $-\beta(a - b)^2$), and so he would not participate.

B will prefer to offer A some less generous contract, one that creates less joint surplus but that gives more of it to B. Since he must persuade A to participate, B cannot extract the extra surplus from a socially-superior contract by reducing all the side payments. So B is concerned not to maximize joint surplus, but to maximize his share of that surplus—a different thing altogether.

[3] This amount, the *expected externality* of A's report a', can be calculated as $\alpha^2\beta(Eb - a')^2$.

To determine just what contract will emerge, we need a theory of the bargaining between A and B. For our purposes of comparison, we will use a theory that gives as efficient an outcome as any: B has all the bargaining power, and makes a take-it-or-leave-it offer to A. Because B does not know a, he cannot simply ask for the efficient value of x and give A just enough money to make him agree (rather than set $x = a$). Instead, he proposes a whole schedule of possible (x, p) pairs among which A can choose; this is like the problem of a discriminating monopolist who doesn't know the demand curve of the customer with whom he is dealing. Maskin and Riley (1984) have analyzed a class of such problems, and show how B's incentive to skimp on side-payments makes his privately optimal scheme one that is not socially optimal.

One can show (see Appendix) that under the contract that B chooses, A's choice x depends on a as shown in Figure 1. For very low values of a, A ignores B's bribes and sets $x = a$. For higher values of a, A accepts a bribe, and sets $x = x^* - \alpha(a_+ - a)$, which is above a but below x^* (except at a_+, where x reaches x^*).

From this we can calculate the ex-ante average payoffs for A and B under this property-rights rule. The algebra is easiest in the case where A always accepts a bribe, as he does provided that C is large enough compared to r and s. In that case (see Appendix), the welfare comparison is ambiguous: *depending on the parameters, the outcome of negotiation may be more or less efficient on average than the bumbling bureaucrat.*

This ambiguous result should make us hesitate to use the Coase theorem to argue for laissez-faire. It is tempting to believe a toned-down version of the "theorem" that claims that private negotiation is a good way to solve externality problems if (1) there is private information; (2) the government cannot effectively use that information; (3) no artificial barriers exist to block voluntary private contracts; and (4) the parties who must negotiate are few and easily identified. Our analysis shows that this conclusion does not generally hold. When there is private information, voluntary private contracts are only imperfectly efficient. So the comparison even with a very bumbling bureaucrat can readily go either way.

Our analysis traces the problem to the effects of private information, which is essential to any interesting decentralization. But private information alone is not enough to cause a problem, as we saw in the discussion of mechanism design (Solomon's schemes). Rather, it seems to be the combination of private information and voluntary participation. Freedom to choose, which is an important safeguard against abuses by central authority, is also a barrier to efficiency.

Does the Coase Theorem Recommend an Institution or Make Institutions Unnecessary?

The Coase theorem is often viewed as a decentralization result that recommends a particular institution: well-defined property rights and voluntary private bargaining over them. As I have shown above, that recommendation is not convincing. The "first-best" decentralization result holds only in special, and rather uninteresting, cases. And the "second-best" result fails the test of our quadratic-uniform example: property rights can easily be less efficient than the bumbling bureaucrat.

An alternative is to see negotiation not as a substitute for other institutions, but as a supplement to them. In this view, private negotiation is a back-up system that may repair efficiency when the main mechanism breaks down. If the back-up system were perfect, there would be no point in evaluating the main system; but it is not. Thus Coase's farmers and beekeepers rely on markets where they can, and negotiate only when for some reason markets fail them. A similar idea is central to the transactions-cost theory of the firm (Coase, 1937; Williamson, 1985; and others), where it is suggested that people create non-market institutions where the market breaks down, and design them not so much to perform well in isolation as to provide a good patch on the overall market system. Arrow (1974) has also argued this point.

This view implies that all economic institutions are better than they seem. Any deficiencies can be repaired, to some extent, by private negotiations. Of course, those repairs will not be perfect, as we have seen; and there may be better repair systems for any particular problem. But negotiation is almost automatically available, and is perhaps relatively unlikely to do harm.

While all institutions are better than they seem, they need not be better to the same degree. One institution may nicely solve problems that would in any case easily be negotiated away, but leave gaping holes that negotiation cannot plug; another may do badly on many problems, but they may be problems that negotiation can readily solve. This suggests a *two-stage* evaluation of institutions. First, see what outcomes the institution by itself will yield; but second, before evaluating its overall efficiency, ask how far its problems will be repaired by negotiation. Since, as we have seen, the efficiency of negotiation depends on the degree of conflict and uncertainty about preferences and the status quo, the Coase theorem does not necessarily mean that we should fire the bureaucrats: rather, it suggests that the way to evaluate them is to ask whether the clumsy compromise that they are prepared to enforce is a good starting point for negotiation, compared, say, to one party's most-preferred outcome. In this model, if both parties agree on a move away from the bureaucrat's compromise, they must both be better off; and so this use of the bureaucrat, to enforce an equitable status quo for bargaining, is more efficient than just letting the bureaucrat decide—hence certainly often more efficient than private property rights of the kind we have analyzed.

I began by noting that the popular simple view of the Coase theorem is a tautology: that if bargaining and negotiation are perfect (that is, produce perfect outcomes) then the outcomes are perfect. Actually, negotiation is far from perfect, even in the simplest situations. (I once tried to bribe some noisy neighbors to be quiet, and the response was a puzzled and angry rebuff.) And it is especially imperfect in the hardest problems—those with private information—where we are most in need of good systems for resolving conflict.

Yet the Coase theorem is much more than simplistic overoptimism or circular reasoning. Organized markets in standardized commodities are not the only institutions for economists to analyze. People can be ingenious in seeking to improve their lot, and even when markets fail some hope remains for cooperation and efficiency. Only if coordination and negotiation problems were trivial would this argument render supererogatory the analysis of institutions: most of the time, most people do what the existing institutions encourage them to do, without asking whether some better contract would be possible. But economists should not forget that people can be creative and can bypass unsatisfactory institutions.

Appendix
The Uniform Quadratic Model

In this appendix, I present the mathematics behind the claims made above about the uniform-quadratic model. I analyze the problem facing person B when he negotiates with A about making side payments to persuade A to set x above a and closer to b. B's choice of a menu of bribes p and choices x obviously implies, and in fact is equivalent to, choosing, for each possible value of a, a payment $p(a)$ and an outcome $x(a)$, with the goal of maximizing the expected value (thinking of a as random) of

$$v(x(a), b) - p(a) = -p(a) - \beta(x(a) - b)^2. \tag{1}$$

In choosing functions $x(\cdot)$ and $p(\cdot)$ to maximize equation (1), B is bound by two sets of constraints. First, since a is not directly observable, it must be the case that A prefers to choose $(x(a), p(a))$ rather than some alternative $(x(a'), p(a'))$ when his true type is a. Formally, for all a and all a', we have

$$p(a) - \alpha(x(a) - a)^2 \geq p(a') - \alpha(x(a') - a)^2. \tag{2}$$

Secondly, we require that A always want to accept B's offer. (This is no loss of generality, since B can set $x(a) = a$ and $p(a) = 0$ if he wants.) Thus,

$$p(a) - \alpha(x(a) - a)^2 \geq 0. \tag{3}$$

The solution of B's optimization problem (1) subject to (2) and (3) is relatively straightforward. Similar problems are treated at some length in Maskin and Riley (1984), for instance. The essence of the solution is to see first that since higher a's always have steeper indifference curves in (x, p) space, we can replace (2) with its first-order version

$$dp/da = 2\alpha(x(a) - a)\, dx/da. \tag{4}$$

Secondly, notice that if (3) binds then (4) cannot, and vice versa. And finally, since higher a makes it more attractive for A to take the payment p and do what B asks, it must be the case that (3) binds for low a (if any) and (4) binds for high a (if any). So, let z be the cutoff value of a where we switch from (3) binding to (4) binding. Then obviously, for $a < z$, $x(a) = a$ and $p(a) = 0$. Next, since z is indifferent between setting $x = z$ (with no payment) and setting $x = x(z)$ and collecting $p(z)$, we have

$$p(z) = \alpha(x(z) - z)^2. \tag{5}$$

Now let $q(a)$ be the shadow price (Lagrange multiplier) on constraint (4) at the point a ($a > z$). Then B chooses the functions $x(\cdot)$ and $p(\cdot)$ over the range $a = z$ to $a = a_+$, as if to maximize the expected value of the Lagrangian

$$-p(a) - \beta(x(a) - b)^2 + q(a)[dp/da - 2\alpha(x(a) - a)\,dx/da]. \tag{6}$$

Applying the calculus of variations to (6) gives us

$$dq/da = -1, \tag{7}$$

$$2\beta(b - x(a)) - 2\alpha q(a)\,dx/da = d/da\{-2\alpha(x(a) - a)q(a)\}. \tag{8}$$

Since q is zero at $a = a_+$, we have $q(a) = a_+ - a$, and then (8) can be solved to give

$$x(a) = 2\alpha a - \alpha a_+ + \beta b = x^*(a) - \alpha(a_+ - a). \tag{9}$$

It remains only to calculate the choice of the cutoff z. By increasing z a little, B loses the benefit of persuading some types of A (near z) to raise x above a. From (9), the choices $x(a)$ of others are not affected. To set against the losses from the types near z, B saves on side payments in two ways: he need not pay those who now set $x = a$; and the entire $p(\cdot)$ schedule above z is lowered slightly. It is straightforward to calculate that the first-order condition on the choice of z implies

$$a_+ - z = \{\beta/(\alpha - \beta)\}(b - a_+). \tag{10}$$

Notice that, as depicted in Figure 1, this is the point where the line (9) meets the line $x = a$. Notice moreover that if (10) cannot be satisfied for any value z in $[a_-, a_+]$, for instance if $\alpha \leq \beta$, or if $\alpha > \beta$ and

$$a_+ - a_- < \{\beta/(\alpha - \beta)\}(b - a_+), \tag{11}$$

then all types of A take the bribe, and $x(a) > a$ for all a. So in symmetric ($\alpha = \beta$) problems, or in asymmetric problems in which C is large compared to the range of a, we can calculate as if (9) held for all a. This makes the calculations considerably simpler. We now proceed to calculate the expected inefficiency resulting from (9), in that simple case.

Since (9) says that $x(a)$ differs from $x^*(a)$ by $\alpha(a_+ - a)$, and since the inefficiency of any rule $x(a, b)$ is the expected squared deviation of $x(a, b)$ from

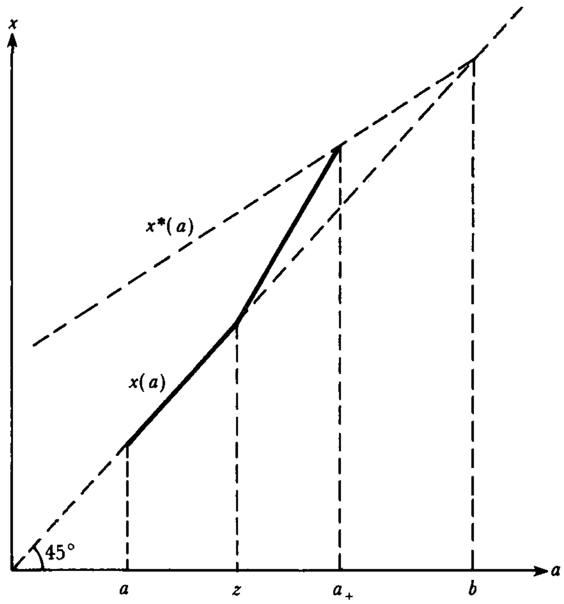

Fig. 1. x as a function of a when A has the right to choose x

$x^*(a, b)$, it follows that the inefficiency of the property-rights rule is just α^2 times the expectation of $(a_+ - a)^2$. Since a is uniformly distributed, this latter expectation is just $(a_+ - a_-)^2/3$, or $4r$. So the inefficiency of giving the right to A and letting B offer side payments is measured by $4\alpha^2 r$. Similarly, of course, the inefficiency of the opposite rule in which B gets the right to choose x and A must offer bribes is $4\beta^2 s$. But the inefficiency of the bumbling bureaucrat is the expectation of $\{\alpha(a - Ea) + \beta(b - Eb)\}^2$, or $\alpha^2 r + \beta^2 s$. So, for many values of (α, β, r, s), including all those with $\alpha^2 r = \beta^2 s$, *the bumbling bureaucrat outperforms both allocations of rights.*

This example was constructed to make this last point. But other interesting observations also follow. For instance, for parameters such that B's optimal scheme always makes A set $x > a$, it is better to give the right to choose x to the party who cares less about it, or whose preferences are more predictable, than to the other party. This runs contrary to the intuition that it should minimize "transactions costs" if the no-negotiation outcome ($x = a$) is "close" to the optimum; rather, we assign rights to the person who would choose the no-negotiation outcome that is further from the optimum x^*.

■ *The author thanks Patrick Bolton, Robert Gibbons, Paul Milgrom, James Mirrlees, Barry Nalebuff, Suzanne Scotchmer, Carl Shapiro, and Joseph Stiglitz for helpful comments.*

References

Arrow, Kenneth, *The Limits of Organization*. New York: Norton, 1974.

d'Aspremont, Claude, and Louis Gerard-Varet, "Incentives and Incomplete Information," *Journal of Public Economics*, 1979, *11*, 25–45.

Calabresi, Guido, "Transaction Costs, Resource Allocation, and Liability Rules," *Journal of Law and Economics*, April 1968, 67–74.

Coase, Ronald, "The Nature of the Firm," *Economica*, 1937, *4*, 386–405.

Coase, Ronald, "The Problem of Social Cost," *Journal of Law and Economics*, 1960, *1*, 1–44.

Farrell, Joseph, "Rigidity Versus License," *American Economic Review*, March 1987, *77*, 195–197.

Friedman, Milton, and Rose Friedman, *Free to Choose*. New York: Harcourt Brace Jovanovich, 1979.

Groves, Theodore, and J. Ledyard, "Optimal Allocation of Public Goods: A Solution to the Free Rider Problem," *Econometrica*, 1977, *45*, 783–810

Hayek, Friedrich, "The Use of Knowledge in Society," *American Economic Review*, September 1945, *35*, 519–530.

Laffont, Jean-Jacques, and Eric Maskin, "The Theory of Incentives: An Overview." In Hildenbrand, Werner, ed., *Advances in Economic Theory*. Cambridge: Cambridge University Press, 1982.

Maskin, Eric, and John Riley, "Monopoly with Incomplete Information," *Rand Journal of Economics*, 1984, *15*, 171–196.

Myerson, Roger, and Mark Satterthwaite, "Efficient Mechanisms for Bilateral Trading," *Journal of Economic Theory*, 1983, *29*, 265–281.

Roth, Alvin, and J. Murnighan, "The Role of Information in Bargaining: An Experimental Study," *Econometrica*, 1982, *50*, 1123–42.

Rubinstein, Ariel, "Perfect Equilibria in a Bargaining Model," *Econometrica*, 1982, *50*, 97–110.

Samuelson, William, "Comments on the Coase Theorem." In Roth, Alvin, ed., *Game Theoretic Models of Bargaining*. New York: Cambridge University Press, 1985, pp. 321–340.

Sonnenschein, Hugo, "The Economics of Incentives: An Introductory Account," Nancy L. Schwartz Memorial Lecture, Northwestern University, 1983.

Sutton, John, "Noncooperative Bargaining Theory: An Introduction," *Review of Economic Studies*, 1986, *53*, 709–724.

Williamson, Oliver, *The Economic Institutions of Capitalism*. New York: The Free Press, 1985.

This article has been cited by:

1. Alessandro Acquisti, Curtis Taylor, Liad Wagman. 2016. The Economics of Privacy. *Journal of Economic Literature* **54**:2, 442-492. [Abstract] [View PDF article] [PDF with links]
2. 2014. Book Reviews. *Journal of Economic Literature* **52**:2, 543-545. [Abstract] [View PDF article] [PDF with links]
3. 2014. Book Reviews. *Journal of Economic Literature* **52**:2, 535-567. [Citation] [View PDF article] [PDF with links]