1972

The Problem of Social Cost Revisited

Donald H. Regan
*University of Michigan Law School, donregan@umich.edu*

Follow this and additional works at: [http://repository.law.umich.edu/articles](http://repository.law.umich.edu/articles)

Part of the [Law and Economics Commons](https://repository.law.umich.edu/) and the [Public Law and Legal Theory Commons](https://repository.law.umich.edu/)

**Recommended Citation**


This Article is brought to you for free and open access by the Faculty Scholarship at University of Michigan Law School Scholarship Repository. It has been accepted for inclusion in Articles by an authorized administrator of University of Michigan Law School Scholarship Repository. For more information, please contact mlaw.repository@umich.edu.
SOME years ago, in a paper entitled "The Problem of Social Cost," Professor Ronald Coase asserted and argued for a proposition which has since acquired the convenient sobriquet "the Coase Theorem." The proposition is: That in a world of perfect competition, perfect information, and zero transaction costs, the allocation of resources in the economy will be efficient and will be unaffected by legal rules regarding the initial impact of costs resulting from externalities. Note that there are two claims being made, which it is well to separate for purposes of discussion. The first claim is that, under the conditions described, some efficient allocation of resources will be achieved, whatever the legal rule. The second claim is that, under those conditions, the same efficient allocation will be achieved, whatever the legal rule. I shall refer to the first claim as the "efficiency" thesis and to the second claim as the "invariance" thesis of the Coase Theorem. I shall argue that neither thesis can be deduced from the traditional assumptions about individual economic behavior which are the foundation of neoclassical price theory and contemporary welfare economics.

Let me be quite clear (or as clear as I can be in advance of presenting my argument)—my primary point is not that the Coase Theorem is false, but that the truth of the Coase Theorem can only be deduced from an assumption or assumptions which differ in kind from the usual assumptions we make about economic agents. The most traditional sort of assumption, and the assumption which is part of every rigorous proof I am familiar with concerning the existence or efficiency of competitive equilibria, is the assumption that each economic agent maximizes something (utility or profit) in the face of prices he takes as given. We depart from this most traditional sort of assumption, and the assumption which is part of every rigorous proof I am familiar with concerning the existence or efficiency of competitive equilibria, is the assumption that each economic agent maximizes something (utility or profit) in the face of prices he takes as given. We depart from this most traditional sort of assumption in the case of the monopolist or the monopsonist—but even the monopolist and the monopsonist operate in an economic environment

2 I should note at some point, and this seems as good a point as any other, that I do not contest the correctness of the Coase Theorem, or its analogue, in certain cases which do not involve a genuine externality at all, such as the case of products liability, with attention focused on accidents to users of the product. (Accidents to "innocent bystanders" are a genuine externality.)
which responds in a particular known way to each possible choice of the agent under consideration. The next step—to the consideration of oligopoly, duopoly, bilateral monopoly, or whatever—takes us into the realm of strategic behavior and the theory of games. My basic thesis in this essay is that the Coase Theorem is a proposition in the theory of games, and not a proposition about traditional markets or competitive equilibrium. My secondary thesis is that the Coase Theorem even as a proposition in the theory of games is open to doubt.

I

Instead of beginning with Coase's explicit argument for the Coase Theorem, let us consider an argument which is implicit in Coase's original paper, but which has been stated most clearly by Calabresi:

The interesting thing about this [Coase's] analysis, however, is that there is no reason whatsoever to limit it to joint cost causers. Thus, if one assumes rationality, no transaction costs, and no legal impediments to bargaining, all misallocations of resources would be fully cured in the market by bargains. Far from being surprising, this statement is tautological, at least if one accepts any of the various classic definitions of misallocation. These ultimately come down to a statement akin to the following: A misallocation exists when there is available a possible reallocation in which all those who would lose from the reallocation could be fully compensated by those who would gain, and, at the end of this compensation process, there would still be some who would be better off than before.

This and other similar definitions of resource misallocation merely mean that there is a misallocation when a situation can be improved by bargains. If people are rational, bargains are costless, and there are no legal impediments to bargains, transactions will \textit{ex hypothesi}s occur to the point where bargains can no longer improve the situation; to the point, in short, of optimal resource allocation. We can, therefore, state as an axiom the proposition that all externalities can be internalized and all misallocations, even those created by legal structures can be remedied by the market, except to the extent that transactions cost money or the structure itself creates some impediments to bargaining.\footnote{Guido Calabresi, \textit{Transactions Costs, Resource Allocation, and Liability Rules—A Comment}, 11 J. Law & Econ. 67, 68 (1968). [Footnotes in original omitted.]}  

I call this the \textit{a priori} argument for the Coase Theorem—\textit{a priori} because it reaches a conclusion about the result of individuals' economic behavior without any model of how individuals behave. What Calabresi is saying may be rendered briefly as follows: If it cost nothing to get all the individuals in an economy together and to get them to reach an agreement about how the economy's resources should be employed, and if all these in-
PROBLEM OF SOCIAL COST REVISITED

individuals were rational, then necessarily, they would agree on a Pareto-efficient allocation.\(^4\)

Whether this statement is a tautology, as Calabresi says it is, depends on the meanings of the terms used. While Calabresi is careful to tell us what he means by “misallocation,” he says nothing about what he means by “rational,” a chameleon among words. In varying contexts the rational man has been taken to be (inter alios): one who maximizes utility (or profit) in the face of given prices and economic conditions; one who maximizes expected utility (or profit) in the face of risk; or one who adopts a maximin strategy if he is playing a two-person constant-sum game. All of these are standard meanings of “rational,” but none of them suffices to make Calabresi’s statement a tautology, because the bargaining situation he envisages is an \(n\)-person, variable-sum game, where \(n\) is likely to be very large indeed. For participants in \(n\)-person, variable-sum games we do not, so far as I know, have any real satisfactory concept of rational behavior. In particular, what we do not have (which we do have in the simpler situations mentioned above) is a model of behavior which seems to be “rational” both for the individual and with regard to the overall outcome for the community.

To illustrate the problem: In a variable-sum game the players have something to gain from cooperation, but there are usually a variety of ways in which the fruits of cooperation can be divided up. Each individual will wish to see not only that the benefits of cooperation are achieved, but that he gets as large a share of the benefits as possible. He will likely be led to threats of non-cooperation as a device to increase his share. Clearly the threats will be ineffective if they are not believed, and it is unlikely that threats will be generally believed unless they are occasionally carried out. But whenever a threat is carried out, the outcome of the game will be sub-optimal, in the Pareto sense.\(^5\) In other words, what appears to be rational individual behavior will occasionally lead to an “irrational” overall outcome.

Two possible ways of avoiding the force of this suggestion come rapidly to mind. One thing Calabresi might do is to define “transaction costs” to

\(^4\) As I observe in the text below, the \textit{a priori} argument would not establish the invariance part of the Coase thesis even if it did establish the efficiency part. Calabresi apparently regards it as establishing both parts because of an implicit assumption that there is a unique efficient allocation.

\(^5\) Support for the suggestion that it is part of the nature of a threat that it results in inefficiency if it is carried out may be found in Thomas C. Schelling, \textit{The Strategy of Conflict} (1960) and R. Nozick, Coercion, in \textit{Philosophy, Science, and Method: Essays in Honor of Ernest Nagel} (Sidney Morgenbesser, Patrick Suppes & Morton White, eds., 1969). Note that what we are suggesting is that the carrying out of a threat results in inefficiency as compared to the original set of possible outcomes. It does not necessarily result in inefficiency within the set of outcomes available after the maker of a threat becomes irreversibly committed to the behavior he has threatened.
include the social cost of individual bargaining tactics. This would be a clever move, but if we are correct in arguing that rational individual bargaining may not always lead to socially optimal outcomes, then this definition would make the assumptions of zero transaction costs and individual rationality inconsistent. Alternatively, Calabresi might define "rational" implicitly by requiring it to mean whatever is necessary to make his claim tautological. In other words, to say that individuals were rational would mean precisely that in a costless negotiation situation, they would come out with a Pareto-efficient agreement.6

This approach has considerable appeal. If we watched the parties to a variable-sum game make threats and counter-threats and finally end up in an inefficient situation, we might be tempted to think at least one of them must have cut off his nose to spite his face. Of course, if we accept the conclusion that rational individual bargaining tactics involve some unavoidable social cost, this is not necessarily true. We might also reflect that this implicit definition is too symmetrical. As far as the definition is concerned, both parties must have cut off their noses to spite their faces. Each has contributed to the failure to reach an optimum. But that impartial judgment has not nearly the same intuitive appeal. Our intuitive reaction is that at least one party must have behaved badly, but not necessarily both. This suggests that we do have some notions of rational (or perhaps reasonable) individual behavior in this kind of situation at an intuitive level. The problem is to formalize those notions and then consider whether rational individual behavior does or does not guarantee collective rationality.7

6 This interpretation of Calabresi might even be suggested by his statement that the efficiency part of the Coase Theorem can be treated as an "axiom." The word "axiom" has at least two ordinary senses, however. In one sense, an axiom is merely an assumption of a deductive system. In another sense, an axiom is an assumption which is supposed to be self-evident. In the text we are considering the possibility of taking the efficiency claim of the Coase Theorem as an axiom in the first sense, and we are arguing against this on the ground that it is certainly not an axiom in the second sense—that is, it is not self-evident. Calabresi, since he views the efficiency claim as tautological and understands it, presumably does regard it as self-evident.

7 There is an assumption which might be useful in some contexts which is not unlike the Calabresi axiom. It is the assumption that the expected magnitude of the inefficiency which will remain after negotiation in any situation decreases as transaction costs decrease. We will not investigate the consequences of this assumption, but we note two points. First, this assumption definitely does not entail the Calabresi axiom. It does not entail that when transaction costs are zero, the expected magnitude of the inefficiency is zero. Second, even this assumption is by no means self-evident. Transaction costs may be such as to make communication not worthwhile or impossible. But to say communication is not worthwhile is to say threats are not worthwhile. If this is true it may be easier for the parties concerned to reach an efficient position which, for one reason or another, is salient, than it would be if communication were possible and there was some prospect of a threat's being effective. This argument closely parallels Schelling's observation that
I do not mean to assert categorically that rational individuals faced with zero transaction costs will fail to reach a Pareto-optimum, although I have suggested what I consider a good reason for thinking this might often be the case. What I do emphasize is that we have no basis for the claim that a Pareto-optimum will be achieved, since we are offered no model of individual behavior. There are some bargaining models we might consider, such as Harsanyi's, but to investigate them would take us considerably afield. I am not aware of any bargaining model which has the features that have made the competitive market model the main tool of price theorists and welfare economists, namely that the existence and efficiency of equilibria can be deduced from assumptions about technology and individual behavior which are fairly plausible on the whole.

So far we have discussed one highly abstract argument for the efficiency thesis of the Coase Theorem. More specific arguments, including Coase's original argument, generally purport to establish the efficiency claim and the invariance claim together, and will be considered in a moment. Since I shall attempt to show that every more specific argument somewhere includes an implicit appeal to the a priori argument which we have just considered, it is worth pausing to note the connection between the a priori argument and the invariance thesis. More specifically, what we should note is that the a priori argument lends no support whatever to the invariance claim. Even if the a priori argument is accepted, it allows us to say only that in a world of zero transaction costs and rational individuals, some efficient allocation of resources will be achieved whatever the legal rules. Will the allocation be the same, whatever the legal rules? This question we cannot answer, on the basis of the a priori argument, since the a priori argument says nothing about what efficient allocation will be achieved in any bargaining situation, and since it gives us no explicit model of individual behavior from which we might hope to find out. If we adopt the a priori argument for the Coase Theorem, part of the Theorem becomes undecidable, since the allocation of resources resulting from any given legal situation is incompletely specified.

Of course, if instead of the a priori argument we were operating with a specific bargaining model, then we might decide the truth or falsity of the invariance claim. It could be suggested that in a plausible bargaining model

inability to send or receive communications may redound to the benefit of a party thus incapacitated. Thomas C. Schelling, supra note 5, passim. The laws against extortion might be interpreted as an effort selectively to increase transaction costs, in order to promote efficiency.

the outcome of any negotiation ought to reflect to some extent our intuitive evaluation of the strength of the parties' bargaining positions; and it could also be suggested that legal rules are an important determinant of the strength of the parties' bargaining positions; in which case we would expect the invariance claim to be false. In other words, one of the criteria for the plausibility of a bargaining model might be that it falsify the invariance thesis (at least if the model is sufficiently sophisticated so that the invariance thesis can be investigated). The main point, though, is that without a model, the \textit{a priori} argument tells us nothing about the invariance claim one way or the other.

II

Now let us shift the focus of our attention to the invariance claim specifically, and let us begin with a simple and obvious argument against the invariance claim. We incorporate Coase's rancher-farmer example and discussion by reference. Even if Coase's argument is good in the short run, and a change of legal rule does not alter the efficient allocation of resources, there will be changes in the long run. If the original legal rule specified that the rancher was liable, and the new legal rule specifies that the farmer is liable, then after the change of rule the rancher will not be paying some damages he was paying before, and the farmer will be paying a bribe he was not paying before. The rancher, then, will be earning a positive profit, and the farmer a negative profit (since by assumption the original position was a perfectly competitive equilibrium). The ordinary consequence of this profit disequilibrium will be the allocation of new resources to ranching, and the exit of resources from farming. In the long run, then, the allocation of resources will change.

The argument I have just made is not novel. It was put forward by Calabresi some years ago. Unfortunately, Calabresi recanted in the later comment on the Coase Theorem from which we have already quoted. His answer to his own original argument was, essentially, that, given Coase's assumptions, "the same type of transactions which cured the short run misallocation would also occur to cure the long run ones." This implicitly assumes that the long-run changes brought about by the alteration of the legal rule are "misallocations," which require to be cured. But on the face of it we have no reason for

\footnote{Specifically, we would expect the legal rules to influence the outcome of a bargaining or arbitration session by influencing the "conflict point," the outcome which prevails if no agreement is reached.}

\footnote{Guido Calabresi, \textit{The Decision for Accidents: An Approach to Nonfault Allocation of Costs}, 78 Harv. L. Rev. 713, 730 n.28 (1965).}

\footnote{Guido Calabresi, \textit{supra} note 3, at 67.}
thinking they are anything of the kind. It is well known that different efficient competitive equilibria correspond to different distributions of wealth. Even granting, then, that the allocation before the change of legal rule was efficient, and that the allocation after the change must also be efficient, there is no reason to assume they must be the same, if the change of legal rule effected a redistribution of wealth, as there is every reason to suppose it did.

Perhaps it would be more enlightening to speak of "distribution of welfare" than "distribution of wealth." If we do that, then it is plain that we should expect a change in legal rule which shifts certain costs from ranchers to farmers to redistribute welfare away from people who consume farm products or supply factors to farms and in favor of people who consume ranch products or supply factors to ranches. Of course, in the usual model of a competitive economy, the "initial distribution of wealth" refers to the initial distribution of stocks of goods, shares in the profits of firms, and ability to supply factors. There is no reference to legal rules. But then, in the usual model of a competitive economy, externalities are not admitted.

So far we have presented the straightforward argument against the invariance claim, and we have pointed out the fallacy of Calabresi's answer to that straightforward argument. A different answer to the straightforward argument is suggested by Professor Nutter in a recent comment. Essentially, Nutter argues that a change of legal rule will not cause a reallocation of resources, but only a redistribution of rents. What Nutter offers by way of argument is simply an arithmetical example, and his example does fall (or may be interpreted to fall) into the one class of cases where his conclusion is plausible—that is, where the change in legal rule might merely redistribute rents. That is the class of cases where ranching and farming both use factors (in this case land) which, when their activity is favored by the legal rule (that is, when their activity is not liable) command rents at least equal to the total damage caused by the externality. In such cases, the cost of damage may come out of the rents earned by factors used in the activity held liable. But this is hardly the general case. In the general case, there may be no factor which is earning any rent at all in the favored activity. All factors may be earning only their opportunity cost of employment in that activity. Plainly, if no factor is earning any rent in the favored activity, then a change of legal rule which makes that activity disfavored cannot merely redistribute rents.

At this point Nutter, or another defender of the invariance thesis, might argue that the factors employed in the favored activity must be earning some rent, on the ground that a situation in which the factors in the favored

activity were earning zero rent could not persist. Suppose farming is the favored activity. Now, if the factors employed in farming are earning zero rent (the argument goes), then the rancher, instead of paying damages, will bribe the factors currently employed in farming to go elsewhere, and will pay out less in this process, since the total bribes need only exceed the total rents, which, being zero, are less than the current damage payment. Therefore the situation will not persist, and an equilibrium will be reached (or so the argument might continue) only when the total damage is less than the rents being earned in the favored activity and the rancher has no further opportunity for profitable bribes.

First, I should like to observe that this argument, like Coase's original argument, depends explicitly on the possibility of bribes, and therefore introduces non-market behavior. We are sliding back in the direction of the a priori argument. Still, the appeal to non-market behavior may not seem objectionable—the model may not seem implausible—in the context of a fairly specific example such as this one. Let us then admit the possibility of non-market behavior, and consider the next step.

The farmer, who suddenly finds that his units of factors have been bribed out from under him by the clever rancher, is not going to sit still. His first thought might be to hire other indistinguishable units of factors, but the rancher will have forestalled that by bribing not merely the units which were originally employed by the farmer but all units of those factors not to work for the farmer. After all, when one can accomplish his purpose with an infinitesimal bribe, he can afford to bribe widely. The farmer's second thought will be better. He will realize that the rancher has turned a tidy profit by bribing his factors away instead of paying damages, and that he (the farmer) can get a share by the simple expedient of threatening to go back into business by offering higher bribes to the required units of factors. To be sure, if he goes back into business on the same scale he will lose money. (Not much, since a bribe which needs only to exceed an infinitesimal competing offer can afford to be quite small. The farmer also has a natural advantage, in that the rancher must bribe all units, while the farmer must bid only for an insignificant fraction of all units.) But he will collect damages from the rancher, and the rancher will lose more than he will. Plainly, therefore, the rancher will see the light and turn over to the farmer some of his saving. How much? Well, we can't say. As a matter of fact, the rancher is in a very touchy position. Just as he had to bribe all the units of those factors which the farmer employed (and not just the particular units originally employed by the farmer), he will now notice that he is open to precisely the same threat the farmer has just made from all potential farmers! We could pursue this line of speculation, but the point is clear enough. The result is at best in-
determinate, and it seems not unlikely the rancher will finally just stay at the original equilibrium, paying out his damages to a farmer whose factors earn zero rent.  

There is a final argument in favor of the Coase Theorem which has been suggested by a number of writers. The argument is essentially that in a world of perfect information and zero transaction costs, there would not be any external effects between or among firms for legal rules to deal with. If there were two or more firms which were failing to internalize between them an externality which connected their production functions (the argument runs), then it would be profitable for these two firms to merge. The externality would be internalized, efficient production would be achieved, and the joint profit would be greater than the sum of the profits of the original firms. The legal rule is left with nothing to operate on, unless it is supposed to govern the distribution of rewards within the firm, which we assume it is not.

The first thing to observe about this argument is that, like Coase's original argument, it assumes successful bargaining and therefore assumes the truth of the Coase Theorem, at least in a limited context. We are offered no explicit model of the process by which appropriate firms are created. It is simply asserted that in a world of perfect information and zero transaction costs, the right organization of the economy into firms would appear. In other words, it is asserted that a bargaining problem among indeterminate numbers of potential entrepreneurs will be "solved" in an efficient fashion. This assertion may seem fairly plausible if we concentrate on externalities which involve only two or three firms before the merger movement, and if we concentrate only on the two or three entrepreneurs associated with those firms. But we have no license thus to limit our vision. For one thing, many external effects operate among many firms. If we allow complete internalization by merger it is likely that we will end up with non-competitive market structures in many ways.

---

13 Actually, once we have opened up the possibility of bribes and threats, we observe that the rancher may be susceptible to threats by potential farmers even when the original farmer is in the market and collecting damages. As long as there is some land around the rancher not being used for farming, any potential farmer is in a position to damage the rancher by going into farming, and therefore in a position to attempt to extort from the rancher a bribe for not going into business. This is true despite the fact that the potential farmer, if he went into business, would lose money. The outcome is, of course, indeterminate. This suggests that one reason we have legal rules against extortion and "spite fences" and so on, is that in addition to being mean-spirited, such devices are incompatible with the market system. Compare supra note 7.

industries. For that matter, if we allow internalization of "pecuniary" externalities as well as technological ones—that is, if we allow firms to merge for the purpose of becoming monopolies, as it will be in their interest to do—then we should expect a non-competitive market structure in every industry. Of course, we could appeal to the Coase Theorem again, and rely on bargains to cure misallocations resulting from monopoly, but now we have lost the defense of our original appeal to the Coase Theorem, which was that we were assuming it only in a limited context, where it was plausible, to prove it in a broader context. There is another difficulty with this appeal to the Coase Theorem in a "narrow context." Even if we have an externality which affects only a small number of actual firms before the merger movement, it may affect many more potential firms. As we observed in connection with the rancher-farmer case, it may not be enough for the rancher to buy off (or merge with) an individual farmer. So long as farming is the favored activity, it may be necessary to include every potential farmer in the deal, since every potential farmer is in a position to inflict damage on the rancher or on the merged rancher-farmer combination. To sum up, the appeal to mergers or "appropriate" firms does not allow us to deduce the Coase Theorem from traditional and plausible assumptions. We need to assume the Coase Theorem in some contexts, and fairly broad contexts at that.

III

The upshot of all this is that Pigou was right after all, in a very important way. Pigou at least recognized that if we wish to rely on the market to produce an efficient result—that is, if we are going to attempt to deduce the efficiency of the result from the traditional assumptions about individual behavior—then some government intervention is required in order to equate marginal private and marginal social cost. On the other hand, Pigou overlooked, or at least did not draw attention to, a very important feature of the problem, which Coase did draw attention to. That feature is that most externalities are, in Coase's words, "reciprocal." Which is to say, it is very likely that the damage done is affected not by the decisions of one individual alone, but by the decisions of many. The reciprocal nature of most externalities means that Pigou considerably underestimated the difficulty of finding regulatory (tax-subsidy) schemes which would guarantee internalization. In the rancher-farmer case, it will not do just to have the rancher pay damages to the farmer, because then the farmer will have no incentive to take into account the effect of his own decisions on the damage. In general to set up an appropriate tax-subsidy scheme might require as much information on the part of the regulating agency as would be required for centralized decision-
making. The market-mechanism-plus-regulation, then, is no certain high road to efficiency. As to invariance, if we do rely on the market-mechanism-plus-regulation, and if we are fortunate enough to achieve an efficient equilibrium, then there is every reason to believe that different regulatory schemes will produce different efficient equilibria.

Finally, a word to the reader who wonders what difference it all makes. After all, even those who assert the Coase Theorem as a deducible proposition about competitive equilibria never seriously claim that the assumption of zero transaction costs is realistic, so they always admit, as Coase himself pointed out, that we need legal rules, and that what the rules are is important. Aside from the simple matter of theoretical correctness, I think my conclusions are important for two reasons. First, the assumption of zero transaction costs, though it is unrealistic in general, might in certain cases, where the externality is localized, be thought no more unrealistic than the other simplifying assumptions economists work with regularly. In such cases, one who believed the Coase Theorem might claim that having a legal rule really was a matter of indifference, whereas if the Coase Theorem is false, or depends on novel and dubious assumptions, this claim would be without foundation. Second, and more important, once we realize that the legal rule affects (or, if we rely on bargaining, presumably affects) the allocation of resources and the distribution of welfare even granting Coase's assumptions, we will not be tempted to speak as if the only function, or even the primary function, of legal rules when there are transaction costs is to help us approach some unique optimum which would exist if only transactions were costless. If there are externalities, then there are decisions to be made about the distribution of welfare even if we assume an "initial distribution of wealth" which is specified in every other respect than the rules concerning externalities.

An illuminating foray into the difficulties of devising an appropriate tax-subsidy scheme is provided by Otto A. Davis & Andrew Whinston, supra note 14.