THE COASE THEOREM

Steven G. Medema
Associate Professor of Economics, University of Colorado at Denver

Richard O. Zerbe, Jr
University of Washington, Graduate School of Public Affairs

Abstract

The Coase theorem has evolved from an illustrative argument in Ronald Coase’s ‘The Problem of Social Cost’ to a centerpiece of the modern law and economics movement. Along the way, the theorem has generated an enormous amount of controversy and discussion, including numerous theoretical attempts at proof and disproof and empirical and experimental analyses of the theorem’s applicability. This chapter surveys the literature surrounding the Coase theorem and presents an outline of the major issues within the Coase theorem debate. In doing so, it attempts to assess the validity of the various challenges to the theorem’s correctness and the implications of those challenges for the theorem’s applicability, which is a separate issue. The analysis presented here illustrates the importance of transaction costs and property rights within the Coasean bargaining process and the need for further research along these lines to flesh out their implications law and economics.

JEL classification: K00

Keywords: Efficiency, Externalities, Invariance, Property Rights

1. Introduction

Ronald Coase’s seminal essay, ‘The Problem of Social Cost’ (1960), is one of the most cited articles in the economics and legal literatures, and much of this attention is owed to a proposition that has come to be known as the Coase Theorem. While the Coase Theorem is by no means the only idea contained within that essay, it has captured the attention and interest of economists and legal scholars as have few other ideas. (For useful treatments of ‘The Problem of Social Cost’ as a whole, see Zerbe, 1976; Schlag, 1986; and Medema, 1996a; for context, see Coase, 1937, 1959.) Coase argued that, from an economic perspective, the goal of the legal system should be to establish a pattern of rights such that economic efficiency is attained. The legal system
affects transactions costs and the goal of such a system is to minimize harm or costs, broadly conceived (Coase, 1960, p. 2). With this in mind Coase (1960, pp. 2-15) demonstrates the importance of transaction costs by considering the nature of bargaining or of contracts that could be struck by using an example of crop damage caused by straying cattle. He noted that negotiations among affected parties would result in an efficient and invariant outcome under the standard assumptions of competitive markets (especially, that the costs of transacting are zero), as long as rights are well-defined. Specifically, it is necessary to know whether the damaging business is liable or not for damage caused since without the establishment of this initial delimitation of rights there can be no market transactions to transfer and recombine them. But the ultimate result (which maximizes the value of production) is independent of the legal position if the pricing system is assumed to work without cost (Coase, 1960, p. 8).

This is as close as Coase comes in his essay to stating what has come to be known as the Coase Theorem.

A. Theorem(s) and Implications

2. Theorem(s)

Although Coase had set forth this idea already in ‘The Federal Communications Commission’ (Coase, 1959, p. 27), the first formal statement of the Coase Theorem did not come until 1966, when George Stigler (1966, p. 113) offered that ‘The Coase theorem ... asserts that under perfect competition private and social costs will be equal’. Subsequently, the Theorem has been stated in numerous ways, including:

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. (Calabresi, 1968, p. 68, emphasis in original)

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. (Regan, 1972, p. 427)

If transaction costs are zero the structure of the law does not matter because efficiency will result in any case (Polinsky, 1974, p. 1665).
if there were (a) no wealth effects on demand, (b) no transaction costs and (c) rights to pollute or control pollution, the allocative solution would be invariant and optimal, regardless of the initial assignment of rights. (Frech, 1979, p. 254)

In a world of zero transaction costs, the allocation of resources will be efficient, and invariant with respect to legal rules of liability, income effects aside (Zerbe, 1980, p. 84).

a change in a liability rule will leave the agents’ production and consumption decisions both unchanged and economically efficient within the following (implicit) framework: (a) two agents to each externality bargain, (b) perfect knowledge of one another’s (convex) production and profit or utility functions, (c) competitive markets, (d) zero transactions costs, (e) costless court system, (f) profit-maximizing producers and expected utility-maximizing consumers, (g) no wealth effects, (h) agents will strike mutually advantageous bargains in the absence of transactions costs. (Hoffman and Spitzer, 1982, p. 73)

when parties can bargain together and settle their disagreements by cooperation, their behavior will be efficient regardless of the underlying rule of law. (Cooter and Ulen, 1988, p. 105)

a change in [law] affects neither the efficiency of contracts nor the distribution of wealth between the parties. (Schwab, 1988, p. 242)

While in many ways similar to one another, these statements of the Theorem contain important differences, many of which are at the heart of the theoretical debates over the Theorem.

Nonetheless, a casual reading of these statements reveals two general claims about the outcomes. The first is that, regardless of how rights are initially assigned, the resulting allocation of resources will be efficient. This proposition - the ‘efficiency hypothesis’ - is reflected in all statements of the Theorem. The second claim, which is not reflected in all statements of the Theorem, is that the final allocation of resources will be invariant under alternative assignments of rights. This is the so-called ‘invariance hypothesis’. The debates over the correctness of the Coase Theorem, and/or its proper form, have turned on both of these hypotheses, and this struggle has been manifest in the current tendency to appeal to two different versions of the Theorem - the ‘strong’ version, which encompasses both the efficiency and the invariance propositions (reflected in the statements of the Theorem by Regan, Frech, Zerbe, and Hoffman and Spitzer, quoted above), and the ‘weak’ version, which encompasses the efficiency proposition alone (reflected in the statements of the Theorem by Calabresi, Polinsky, and Cooter and Ulen, quoted above).
3. Implications

For economists - Coase’s target audience (Coase, 1988a, 1993) - the implication of the Theorem is that if remedies are considered in the unrealistic world in which competitive markets are normally considered, a world of zero transactions costs, the Pigouvian remedies said to be necessary for an efficient resolution of externality problems are not, in fact, necessary. All that is needed is a common law or statutory rule which assigns rights over the externality to one party or another. The market/pricing mechanism will then function in the same way as it does for ordinary goods and services over which rights are clearly defined. Furthermore, if rights are well-defined, the observed situation will be efficient (the parties having taken all Pareto-improving steps) and any further intervention (for example, Pigouvian remedies) will make matters worse rather than better.

If the implications for the economics of externalities are heretical, those for law are downright perverse, for the Theorem tells us that the form of legal rules does not matter - only their presence or absence. Thus, we will have the same amount of pollution (and thus clean air or water) and of outputs associated with the generation of pollution regardless of whether polluters or the victims of pollution are made liable for pollution damage. The same amount of effort will be devoted to precaution against causing tortuous injury regardless of whether injurers or victims are liable for harm caused. The structure of law pertaining to breach of contract will have no impact on the allocation of resources through the contracting process. Attempts by judges to engage in social engineering from the bench will be fruitless, apart from distributional (as opposed to allocational) effects. Assuming that rights are alienable, the allocation of resources will be the same regardless of the rule of the law, and that allocation will be efficient. More generally, it is a matter of indifference whether courts impose property rules or liability rules (Calabresi and Melamed, 1972), and the entire issue of adherence to precedent becomes a moot point in terms of its effect on the allocation of resources.

B. Is the Coase Theorem Correct?

The Theorem has never been formally proved. Arguments regarding its correctness or incorrectness generally consist of attempts to demonstrate that it does or does not hold in a particular context or under a certain set of assumptions. Of particular import here is the framework within which the reallocations of rights contemplated by the Theorem are assumed to take place. Two basic frameworks can be identified: the ‘quasi-competitive’ framework, within which all relevant markets are assumed to be perfectly competitive and
agents operate more or less along competitive lines in externality negotiations, and the ‘game-theoretic’ framework, within which there exists the potential for strategic behavior among affected parties. We examine each of these in turn.

4. The Quasi-Competitive Framework

Most of the analysis of the Theorem has taken place within the quasi-competitive framework, under which are two different types of treatment. First, there are discussions of small numbers externality negotiations within a more broad competitive context of full information, no strategic behavior, agents operating within competitive markets, and so on, with the result that parties strike mutually-beneficial bargains when they are available. This is analogous to the standard Edgeworth box analysis and is the environment contemplated by Coase in ‘The Problem of Social Cost’. The second type of treatment actually assumes competitive markets in externality rights and analyzes the Theorem on that basis. In the latter case, the first optimality theorem of welfare economics suggests that the Coase Theorem is correct (Arrow, 1969).

4.1 Rents

One of the earliest but merely technical arguments raised against the Theorem is that it cannot hold under perfectly competitive conditions in the long run because it presupposes rents that may not exist. To consider the objection suppose that both the polluter and the victim are in a zero-profit, long-run equilibrium position. (For expositional simplicity, in the following discussion we will refer to the parties as ‘polluters’ and ‘victims’. The analysis, of course, generalizes to all manner of externalities.) Then, the assignment of liability to the polluter will force the polluter to cease operations, since it lacks the resources with which to make liability payments to the victim. Similarly, if the victim is made liable, it will exit the market because it lacks the resources to bribe the polluter to reduce its harmful activity. Thus, it is argued, the Coase Theorem will only hold in the presence of ‘non-transferable resources giving rise to Ricardian rents’ (Wellisz, 1964, p. 351). It has been further argued that even the prior existence of rents would not ensure the validity of the Coase Theorem: the rents must be sufficient to support the externality - that is, sufficient to allow the polluter to pay damages (if he is liable) or the victim to pay the bribe (if the victim is liable); otherwise, a change in the direction of liability will cause the party bearing the cost of the externality to exit the industry (Tybout, 1972; Shapiro, 1974). Let us analyze these claims.

Suppose that the polluter is not earning rents and that he is liable for damage caused. Then, the polluter will be forced to go out of business in the long run, as he does not have the resources necessary to pay damages. This
result is efficient since the externality existed inefficiently in the first place: the polluter was able to inflict damages on the victim only because he did not bear the full social cost of his actions. If the polluter is not liable for damages, the victims would be willing to offer a bribe up to the amount of damage to induce the polluter to cease operations, a bribe which the polluter would be willing to accept since it was not earning any rents in the first place. Thus, whether the polluter or the victim is earning no rents, we get the same, efficient, result regardless of the assignment of rights.

If neither party is earning rents sufficient to support the externality, the externality would not exist in the first place; the appearance of the externality would immediately drive the victim out of business and the externality would cease to exist. (See Wellisz, 1964, p. 350; Crain, Saurman, and Tollison, 1978 and Zerbe, 1980, pp. 89-90; 99.) Thus rents must exist for negotiation over rights to even be in the realm of possibility; that is, they are prior to the Coase Theorem analysis. But once they are satisfied, an efficient and invariant result will obtain.

The one case in which rents are not necessary is for externalities that are industry-wide in their emission and public bads in their effects (so that all firms in an industry are affected). Here, assuming that all firms are harmed equally (in terms of the effect on marginal and average costs), the externality acts in a manner analogous to an increase in input prices, causing an increase in marginal and average cost, a reduction in supply, and an increase in market price. The result will be invariant under alternative rules of liability.

4.2 Entry in the Long Run
One of the most discussed challenges to the Theorem concerns the effect of liability or bribe payments on entry into markets. If polluters are made liable for damages, the flow of liability payments into the victim industry will increase the rate of return in that industry. If one assumes that firms entering the market are also eligible for compensation, then entry will occur in the long run, leading to an increase in the output of the victim industry. When victims are liable, in contrast, the flow of bribe payments from victims to polluters raises the rate of return in the polluting industry, leading to entry into that industry and a corresponding increase in output. The arguments here are two: first, that the invariance proposition fails to hold because of the disparate entry effects of alternative legal rules; second, the efficiency proposition fails to hold because when polluters are liable the bribe-induced entry will result in too much victim output, relative to what is optimal, while when victims are liable the bribe-induced entry will result in too much polluter output, relative to what is optimal. (See Calabresi, 1965; Bramhall and Mills, 1966; Tybout, 1972; Baumol, 1972; Schulze and d’Arge, 1974; and Frech, 1979.)

The inefficiency issue is easily disposed of. First, if transaction costs are zero, agents are rational and there are no legal impediments to bargaining, then
The Coase Theorem

The Coase Theorem explains that the long-run inefficiency will be cured through the same type of bargaining transactions that were employed to resolve the short-run inefficiencies caused by the externality (Calabresi, 1968, p. 67). That is, available gains from exchange will be exploited in the long run just as they are in the short run. Second, following Nutter (1968), any long-run misallocation will be cured by a single owner who will enter the market in order to exploit the potential for gain. Third, the above argument assumes the existence of a single efficient long-run equilibrium point (and thus that efficiency implies invariance), when, in fact, such need not exist. Here, the long-run equilibria are both efficient and thus the subsequent corrective negotiations or entrepreneurial actions are unnecessary. That is, long-run entry effects do not invalidate the efficiency argument.

The more difficult issue is that of invariance. While the above argument against invariance would appear to be straightforward, consider the following counter-argument. Suppose that ranchers are liable for damage done by their cattle. The flow of liability payments will then be capitalized into the value of farmland that adjoins ranching property and there will be no incentive for entry into farming in order to secure the bribe. In analogous fashion, any bribes that result from farmer liability will be capitalized into the value of ranchland that adjoins farms and there will be no incentive to enter ranching. Given this, the long-run entry effects that are said to invalidate the invariance proposition will not occur (Demsetz, 1972a; Frech, 1979).

The key to distinguishing between these competing claims regarding invariance has been provided by Holderness (1989), who pointed out that invariance turns on the issue of whether rights are assigned to open or closed classes of individuals or entities. An open class is defined as one into which entry is unrestricted, while a closed class is one which can be entered only if the right is purchased from a current class member (see also Demsetz, 1972b, 229-231). Consider first the assignment of rights within closed classes. Landowners constitute a closed class since one can become a landowner only by purchasing the land and the attendant bundle of rights from a current landowner. The assignment of rights to one class of owners creates at once a windfall gain for those having the right and a windfall loss for those not having it. However, in a competitive system these windfall gains and losses are immediately capitalized into the value of the land so that both types of land yield a normal rate of return. Since the rate of return for each of these types of land is unaffected by the assignment of rights, there are no incentives for entry or exit. Thus, the invariance proposition holds for closed classes (Holderness, 1989, pp. 183-184).

The invariance claim does not hold for open classes, however. Holderness' separation of open from closed classes calls attention since to a broad category of spurious objections all based on incomplete property right specification. Here, those who are not parties to the lawsuit through which the
initial assignment of rights is generated can acquire that right costlessly merely by entering and this valuable right will not be capitalized into the price of any resource. Entry will indeed result (Holderness, 1989, p. 185). Similarly, the absence of a right reduces the returns to that activity, thereby inducing exit. The asymmetric entry/exit effects across alternative assignments of rights will thus result in different long-run outputs under alternative assignments of rights, thereby negating the invariance proposition.

This distinction illuminates the divergent results obtained by many of those offering support for or claiming to refute the invariance proposition. Those who have found the invariance proposition to be valid in this (the entry issue) context have either explicitly or implicitly assumed or worked with examples constituting closed classes. On the other hand, those finding against invariance have analyzed the problem in open-class contexts - primarily situations with two industries where entry is possible. The invariance proposition is applicable to closed class situations, such as externalities affecting land values, but is inapplicable to tort situations, such as accident law, where there is free entry into one or both classes and to assignments of rights which cover all (current and future) entrants into an industry.

The open class case, however, would seem to violate an underlying assumption of the Coase Theorem - fully-specified property rights. That is, rights in open classes are not delimited to the extent necessary to make market transactions possible; potential entrants are able to secure a valuable right without paying for it. This is consistent with Barzel’s (1989, p. 2) definition of property rights as ‘the powers to consume, obtain income from and alienate ... assets’ and Allen’s (1995, p. 2) definition of an ‘economic property right’ as ‘one’s ability, without penalty, to exercise a choice over a good, service, or person’ (emphasis in original). In fact, the assumption of zero transaction costs is said by some to mean that rights are fully specified (see Cheung, 1992; Allen, 1991, 1995) and the discussion in section 6, below). As such, the issue of incompletely specified rights, or open classes, goes to the issue of relevance rather than correctness.

4.3 Separable v. Non-separable Cost Functions

Another class of interesting objections is based on a failure to consider fully contract or merger possibilities. It has been argued, for example, that the validity of the efficiency (and invariance) proposition will turn on (in addition to other previously-recognized problems) the form of the cost function of the victim firm. It is well established that, if the victim’s cost function is additively separable (that is, if \( C_n = C(q_n, q_0) = C(q_n) + C(q_0) \)), then the Coase Theorem holds - the outcome is both efficient and invariant under alternative assignments of rights (see Gifford and Stone, 1973; Marchand and Russell, 1973). Suppose instead, however, that the victim’s (B’s) costs of production are dependent upon both its own output and the output of the polluting firm (A),
so that \( C_B = C(q_A, q_B) \), where the pollution damage owing to firm A’s output increases the costs of firm B.

With a non-separable cost function there is neither efficiency nor invariance. With a non-separable cost function, the level of pollution damage to B is a function not just of A’s output, but of B’s output as well and thus a given level of output by A causes B more harm (that is, causes a greater increase in B’s costs) the more output B produces. In such a situation, the victim can and does contribute to its own damage without having to bear the cost, since it is fully compensated for all damage. As a result, the victim has no incentive to mitigate damages and produces an inefficiently high level of output. Moreover, the damage liability associated with this imposes a higher than optimal cost on the polluter, causing it to restrict its output below the optimal level (Marchand and Russell, 1973, pp. 613-615).

However, if both activities are controlled by a single owner, the result will be efficient and invariant regardless of the initial assignment of rights and irrespective of whether the victim’s cost function is separable or non-separable (Marchand and Russell, 1973, pp. 614-616). Moreover, one can imagine a contract between owners that mimics the effect of single ownership assuming costs of monitoring and negotiation are zero. This demonstration is sufficient to negate the nonseparability critique, since the inefficiency contemplated will be exploited through merger, which can be achieved costlessly, or by an entrepreneur (see, for example, Nutter, 1968; Coelho, 1975 and Zerbe, 1980, pp. 87-88). Marchand and Russell (1975) have responded to this criticism by invoking difficulties in carrying out a merger - that is, by introducing transaction costs. But introducing transaction costs is no argument against the correctness of the Coase Theorem. A further argument that can be raised against the nonseparabilities critique is that it violates the assumption of fully-specified property rights (at least in the sense of Allen, 1991, 1995), since the victim is able to procure revenues from the polluter without giving up anything in return.

4.4 Non-Convexities at the Negotiation Starting Point
Perhaps the one seemingly insurmountable criticism of the invariance proposition comes from the recognition by Starrett (1972) that externalities will cause nonconvexities to exist in the production sets of victim firms. (See also Shapiro, 1974, 1977, 1978; Vogel, 1987.) The argument applies equally to consumption sets and thus to externalities to which consumers are party. This objection is similar to that of non-separable cost functions in calling attention to a contacting problem. This is an interesting objection because it points to the importance of information costs. In graphical terms, rather than generating a convex production set, such as in Panel a of Figure 1, the externality causes nonconvexities of the form illustrated in Panel b.
Here, $q_B$ is the output of firm B, the victim and $Z$ is the level of the externality. The analysis turns on the effect of increasing pollution damage on the victim’s output. Panel (b) illustrates a situation of increasing marginal damage from pollution (given by the reduction in the victim’s output due to the externality) with the recognition that, beyond some point ($Z_0$), marginal damage will be zero (Starrett, 1972, pp. 189-190).

The import of this for the Coase Theorem is as follows. Suppose that the polluter’s (A’s) profit-maximizing level of pollution in the absence of the legal rule is some level $Z_1 > Z_0$. With this level of pollution, B will produce no output. If the A has the right to pollute, the point $Z = Z_1$, $q_B = 0$ is the starting point for negotiation over the level of pollution. The minimum payment that the polluter is willing to accept to reduce pollution is the reduction in profits that would accompany the reduction in pollution. However, at (and around) $Z_1$, there is no benefit to the victim from a one unit reduction in pollution; the victim’s output ($q_B$) would remain at zero with this one unit reduction in pollution. Thus, the victim would not be willing to offer a bribe payment to induce the polluter to reduce its pollution by one unit (to $Z_1$! 1 units) - it is a cost with no attending benefit. Thus, the equilibrium when the polluter has the right to pollute will be at a pollution level $Z = Z_1$ and an output level for the victim of $q_B = 0$, a result which is due to the nonconvexity.

On the other hand, if the victim has the right to be free from pollution, the baseline from which negotiation begins is $Z = 0$. The victim will be willing to accept any bribe to allow positive levels of the pollution if the bribe is in excess of the lost profits due to pollution damage (or reduction in $q_B$). Thus, the parties will be able to negotiate to an efficient result along standard Coasean lines, but the final result will not (except by accident) be $Z = Z_1$ and $q_B = 0$. Thus, the
The Coase Theorem

The invariance proposition does not hold in the presence of nonconvexities if the starting point for negotiation when the polluter has the right to pollute falls within the non-convex region of the victim’s production set. However, if the initial level of pollution is some level \( Z \# Z_0 \), the invariance proposition holds. (See Cooter, 1980 for a discussion of how legal rules can be specified to circumvent the nonconvexity problem).

The force of this critique is sufficiently powerful that the editors of the Journal of Economic Theory said that Starrett’s demonstration that nonconvexities are inherent in externality problems ‘destroy[s] the validity of the Coase Theorem’ (Shapiro, 1977, p. 222). However, this judgment fails to understand the true nature of a fully transaction costs world. There is a Pareto-better point available, but the nonconvexity means that the parties will fail to reach it because the immediate marginal adjustments are not Pareto-better. Essentially, the point is that the victim will not be able to spend $15 for a change that will make it $20 better off because the first step along this path would involve spending a dollar to get a zero improvement in welfare and the victim, not being willing to take this first step, will never know that better things are on the horizon. The problem here is one of information: it is certainly the case that the victim would take this first, welfare-reducing step if it was certain that, in the end, it would be better off. If both victim and polluter knew of the existence of a superior position they could also merge to achieve it. Thus, the nonconvexities argument introduces imperfect information into the model. If, as most have maintained, information costs violate the zero transaction costs assumption of the Coase Theorem, fundamental non-convexities do not produce inefficiency. The non-convexity critique of the Theorem points to the importance of information costs, but, if these are considered as part of transactions costs, it does not point to the incorrectness of the Theorem itself.

4.5 Income, Taste and Preference Effects

An additional complication first hinted at by Buchanan and Stubblebine (1962) and Turvey (1963) and later more explicitly elaborated by Dolbear (1967) and Mishan (1965, 1967, 1971), within the economics literature, and by Kelman (1979) within the legal literature - is introduced when one or both parties to the externality are consumers, for, at this point, we are forced to take into account effects on demands that attend alternative assignments of rights. These effects arise first, from differences in tastes between parties and, second, from differences between the willingness to pay (WTP) and willingness to accept (WTA).

Differences in Tastes If tastes are different (that is, as long as indifference curves are not homothetic) a change in the income distribution will affect the pattern of demand and therefore the pattern of resource allocation, though not
its efficiency. That is, alternative assignments of rights can have differential effects on the structure of demands for consumer goods. Since different assignments of rights result in different distributions of income, the composition of demands - and hence equilibrium prices and quantities across markets - will vary with alternative assignments of rights. For example, suppose that an economy produces only beef and fish and that fertilizer runoff used to produce grass, an input into beef production reduces fish production. Those that produce beef prefer to eat beef and those that produce fish prefer fish. A change in the liability rule from one in which beef producers are liable to one in which they are not will increase the income of beef producers and reduce the income of fish producers. This will then increase the demand for beef and reduce the demand for fish; thus, the relative production of beef and fish will not be invariant to the liability rule. This objection to the strong version of the Theorem is well recognized. But, we shall see that even this objection presupposes incomplete property rights.

**Differences Between WTP and WTA** A change in the law can change the sense of ownership and thus change the measure of value from the WTP to WTA or vice versa. Following Willig (1976), economists have tended to assume that any differences between willingness to accept (WTA) and willingness to pay (WTP) owing to a price change are small. This is now recognized as untrue in important cases. For environmental goods, researchers have demonstrated repeatedly that WTA questionnaires generate values from two to nineteen times greater than those elicited by WTP questions (Levy and Friedman, 1994, p. 495, n. 6; Hoffman and Spitzer, 1993, pp. 69-85). There are three reasons for the difference: income effects, substitution possibilities and loss aversion. (Hoffman and Spitzer, 1993) present an excellent survey of the WTA v. WTP issue, much of the evidence regarding which comes from the experimental literature. The present discussion touches on what we believe to be the most significant of these arguments.)

Let us first consider the implications of income effects. If most people experience diminishing marginal utility of income, the utility loss resulting from a reduction in income of a certain amount is greater than the utility gain associated with an increase in income of the same amount. Thus, if individuals bargain over utility, rather than over wealth *per se*, we would expect to see differences between WTA and WTP and thus negotiated solutions that vary with the initial assignment of rights (Hovenkamp, 1990). This is an income effect. For example (and assuming that A, the polluter, is a firm, so that wealth effects are irrelevant for it), if B (an individual) has the right to be free from pollution, then the amount of pollution generated will be a function of the payment that he is willing to accept to avoid pollution. If B does not have the right to be free from pollution (that is, A has the right to pollute), then the amount of pollution generated will be a function of the amount that B is willing to pay to avoid pollution. Since the maximum amount a person is willing to pay
to avoid damages is a function of his budget constraint, while there is no such
constraint on the amount that the individual is willing to accept, we will see (as
long clean air is a normal good) a difference in the amount of pollution and
thus in pollution-related output, depending on the initial assignment of rights
(Mishan, 1971, p. 19). Thus, the invariance proposition will hold only when
income effects are not present or when all relevant income elasticities of
demand are zero.

The force of the income effects critique has been reflected in one of the few
major modifications of the standard structure of the Coase Theorem: the
addition of the assumption of no income effects or the qualifier ‘income effects
aside’, as reflected in several of the statements of the Coase Theorem set out
above. The use by some of the weak rather than the strong version of the
Theorem is also attributable largely to the role that income effects play in
negating invariance.

Consider next the issue of substitution possibilities. Recently, Hanemann
(1991) showed that the poorer the substitutes for the good, the greater the
divergence between the WTP and the WTA. Put another way, the divergence
will be greater the more unique the good. The substantial divergence between
WTP and WTA for unique goods arises in part from the fact that many of these
goods have no close substitutes. Thus, for most people, the WTA to allow
degradation of the Grand Canyon will be much greater than the WTP to
prevent degradation for most people. A change in the law that results in a
difference from a WTP to a WTA criteria will have dramatic effects in the
measure of value for the good. Zerbe (1998a) and Cohen and Knetsch (1992)
have argued that the correct measure of damages as between the WTP and
WTA is, however, a function of the psychological reference point which may
not correspond with the legal reference point. For example, the Ellickson
(1986) study of the response to differing range laws shows such a difference.
In one half of Shasta County, California, open range was the legal rule and in
the other half closed range was the rule. Yet in both parts of the County, in
spite of opposite assignments of liability, people expected and provided similar
remedies. Cattle owners took responsibility for the damages in all cases and this
responsibility existed for many years and was enforced through social norms.
The reference state was one of crops not being damaged by straying cattle.

A further argument against the invariance proposition comes from the
influence that alternative assignments of rights may have on WTA v. WTP
through effects on consumer tastes and preferences. Here, the assertion is that
consumer tastes and preferences are not wholly exogenous to the structure of
legal rules but are influenced by them (Zerbe and McCurdy, 1996). Prospect
tory theory posits that individuals have a value (rather than utility) function which
is convex for gains and concave for losses and that the degree of concavity is
greater than the degree of convexity, so that losses of a given size are felt more
acutely than gains of that same size (Kahneman and Tversky, 1979). The link
between this idea and the WTA v. WTP argument is that the offer of money to
relinquish a right would be treated as a loss, whereas the purchase of a right would be regarded as a gain. Since losses count more than gains, the minimum amount that an individual would be willing to accept to relinquish a right will exceed the amount that he is willing to pay to acquire it (Kahneman, Knetsch and Thaler, 1990). One reason for this loss aversion is suggested by what Thaler (1980) has called an ‘endowment effect’, owing to the fact that people value ‘received income’ more highly than ‘opportunity income’. Because of this, people will be willing to forego more opportunity income to retain a right than they would spend in received income to acquire it and thus WTA will exceed WTP (Thaler, 1980; Kelman, 1979). Kelman attempts to apply this argument to the producer side as well, arguing that producers may value realized and opportunity income differently. However, doing so would contradict the assumption of profit maximization. Moreover, as Spitzer and Hoffman (1980, p. 1210) point out, a profit-maximizing entrepreneur could (and, following Nutter, 1968, would) arbitrage this difference, thus generating the outcome implied by the Coase theorem.

How important these forces are in creating a wedge between the WTP and the WTA is as yet uncertain. At this point, however, it seems reasonable to say that income and substitution effects and loss aversion are sufficient to invalidate the invariance, although not the efficiency, claim of the theorem.

The above possible exceptions to the Coase Theorem represent important cases, but do they really represent exceptions to the Coase Theorem? All of the objections to the Coase Theorem that rest on consumer preferences rest on a change in the distribution of wealth. Yet, these arguments at base reflect property rights that are not fully specified or are inefficiently specified. These conditions then violate the Coase Theorem assumptions that property rights are fully specified.

Consider first a change in the rule of liability. If property rights are fully defined (in the sense of complete ownership), this alteration of liability cannot take place without compensation; if it does, the right was not fully defined in the first place, in violation of the Theorem’s assumptions. Thus, owing to the compensation, the distribution of wealth will be unaffected (Allen, 1995, p. 10). Allen’s argument applies even to the income effects qualification. Of course, this rebuttal does not go to the case where non-existent rights are subsequently defined. However, in a world of zero transaction costs the definition of rights would be perfectly anticipated and thus reflected in resource values (Allen, 1995, pp. 10-11). In sum, when alternative assignments of rights influence the distribution of income and wealth, it must be the case that rights are less than fully defined and/or that transaction costs are positive. Indeed, Allen (1991, 1995) has argued that fully-defined rights and zero transaction costs are really the same thing (see the discussion in Section 6, below).
Perhaps the most intriguing case is one suggested by the recent literature in which legal ownership is different from psychological ownership (Zerbe, 1998a, 1998b). Evidence suggests that a sense of ownership attends certain environmental goods even if there is no individual ownership. A decision to cut down the last remaining stand of privately owned redwood trees, the Headwater Grove, may create a sense of loss among some that are non-owners of the grove. This loss is correctly measured by the WTA. If, however, a decision to measure the value of the grove is based on legal ownership the value to the public will be based on the WTP for preservation. Since for a normal good the WTP will be less than the WTA, the grove may be cut when it should not. Property rights are fully specified in this example so that it would appear to violate the strong version of the Theorem.

However, Zerbe (1998a, 1998b) has noted that, although property rights in this case are fully specified, they are inefficiently specified. He argues that 'efficiency requires that the legal measure of property and damage correspond to psychological reference points'. (Posner’s rule (1992, p. 52) for the allocation of rights is a subsidiary of this theorem. This rule is that where one class of claimants values the right more than other classes, efficiency requires that the right should go to the claimants that value it the most.) If there is not a correspondence between psychological and legal property rights, the use of WTP and WTA based on legal criteria can impose net losses. Imagine that a party, George, believes he owns a right or a property, M, and that another party, Ronald, also believes that George owns property M. They discover that the law, however, holds that party George, not Ronald, owns M. Ronald suffers a loss of M psychologically and therefore economically, while George gains M. Since losses are, on the average, worth more than equivalent gains (due to income, substitution effects and loss aversion), on the average George will gain less than what Ronald loses. This is perfectly general, so that the application of law to affect a legal ownership different from psychological ownership must, on average, impose net losses. (This is true as long as Ronald and George may be regarded as equivalent in the sense that on average one does not have a greater income than the other or does not differ in some other relevant characteristic. Underlying this proof is the notion that we cannot speak of it being efficient to change preferences to be in accord with the law since this violates the proper context for benefit cost analysis - which requires that preferences be taken as they lie - and the very concept of efficiency. In any event benefit cost cannot evaluate the advantages of a change in preferences since this does not take preferences as they lie.) Similarly, if one class of claimant psychologically possesses property so that its removal is felt as a psychological loss, as compared with a rival claimant who has a lesser psychological claim or no claim, efficiency requires that the law grant the right to the psychological possessor. But, in a zero transactions cost world this sort of inefficiency would
not arise since the law would be made to correspond with the psychological sense of ownership.

5. The Game-Theoretic Framework

While the vast majority of the literature debating the validity of the Coase Theorem employs the quasi-competitive framework, a number of commentators have addressed the Theorem from a game-theoretic bargaining perspective, arguing that the quasi-competitive framework is not appropriate or relevant for Coase Theorem-like bargains over rights owing to the small number of parties contemplated and the potential for strategic behavior. (See, for example, Davis and Whinston, 1962; Samuelson, 1966, p. 1141; Shoup 1971, p. 310; Regan 1972, p. 428; Cooter, 1982, pp. 16-17.) By placing the Theorem in a small numbers context but yet ignoring the potential for strategic behavior, it is presumed that the contemplated agreements ‘can and will be reached because it is in the joint interest of the parties to do so’ (Samuelson, 1985, p. 322). Yet, this fails to consider the possibility that what is rational for the group may not be rational for the individual and constitutes, in essence, an ‘a priori argument’ for the Theorem (Regan, 1972, pp. 429-431).

Several commentators (for example, Davis and Whinston, 1965; Arrow, 1979; Aivazian and Callen, 1981; Samuelson, 1985 and Aivazian, Callen and Lipnowski, 1987) have suggested that the Coase Theorem, as envisioned by its proponents at least, lends itself quite naturally to the theory of cooperative games. It can be demonstrated that the Coase Theorem will always hold in a two-person cooperative game. However, this result is not particularly comforting since, by setting the problem in the context of a two-person cooperative game, efficiency is assured by definition (although there is no guarantee of invariance), making this more along the lines of an illustration of the Coase Theorem rather than a proof (Schweizer, 1988, pp. 246, 254). In fact, much of the quasi-competitive literature (especially the two-person analysis) can, without too much injustice, be described as cooperative game analysis. The situation is complicated when the cooperative game involves more than two players. While it has been suggested that the Theorem may not hold in such a context (Aivazian and Callen, 1981), this claim has been shown to be incorrect (Coase, 1981; De Bornier, 1986).

5.1 Noncooperative Game Theory

However, the most interesting and potentially most damaging, game-theoretic analysis of the Theorem has involved the use of noncooperative game theory. If parties have full information about each other’s utility (or profit or production and cost) functions, the Coase Theorem will hold in a
noncooperative setting. The initial assignment of rights establishes the utility level of each player in the absence of further reallocations of resources and there are assumed to exist reallocations of resources which are efficiency enhancing, in the sense that the utility of one player can be increased without reducing the utility of the other player. However, neither party will agree to an alteration in the allocation of resources unless that reallocation increases its utility. The question, from the perspective of the noncooperative game, is whether there exists a sequence of permitted moves which will generate an efficient (Nash) equilibrium. Suppose that A is given the right to pollute. Then the victim, B, has an incentive to offer an alternative allocation of resources (for example, a combination of a bribe paid from B to A and a reduced level of pollution by A) to A, but B knows that A will accept this offer only if A’s utility is increased under the new allocation. Since B knows A’s utility function, he can determine the range of allocations sufficient to garner A’s acceptance. And, since B’s utility is higher with reallocation than without, he will not offer an allocation that A would reject, choosing instead that allocation from the group A will accept which maximizes his own utility. The resulting equilibrium is Pareto efficient, since, given the utility level of one player, the other player’s utility is maximized. As Arrow (1979, p. 29) points out, however, this is not the competitive equilibrium. The same reasoning applies to the situation where the victim is given the right to be free from pollution (Arrow, 1979, pp. 27-29; Schweizer, 1988). In fact, it is not necessary for each party to enter the bargaining process with full information, only that each party perceives that there are net gains to it from providing full information during the negotiation process and thus will reveal such information during that process (Saraydar, 1983, p. 603, n. 12).

The problem, numerous commentators have pointed out, is that it is unlikely that the players will know each others’ respective utility (or profit, production or cost) functions. This has a number of implications.

(a) Implication One: The Baseline Problem When victims are liable, the firm can influence the level of the bribe that it receives by making an upward adjustment in pollution emission at the time that the baseline level of pollution (against which subsidized/bribe-induced pollution reductions will be measured) is set or by choosing not to take cost-justified precautions. That is, the level of pollution on which the bribe is based may differ from the level of pollution that would have been emitted if the polluter was liable. The source of this incentive is the inability of the victim to ascertain with certainty the true baseline level of pollution. Moreover, disagreements over the baseline level of pollution may result in the failure to consummate bargains when the victim is liable. Conversely, if polluters are known to be liable for damages, then, in the absence of full information about actual damages and measures taken by the victim to
mitigate damages, the victim’s moral hazard will result in too few resources being devoted to precaution/mitigation by the victim and too many resources being devoted to abatement by the polluter (Kamien, Schwartz and Dolbear, 1966; Tybout, 1972; Harris, 1990). With endogenous liability assignment, however, the moral hazard problems disappear, since a party does not know with certainty whether or not it will be forced to bear the costs of the externality. Thus each party will act efficiently to minimize expected costs by engaging in the appropriate level of precaution/preventive activity (Harris, 1990, pp. 701-702).

(b) Implication Two: Extortion Imperfect information raises the problem of extortion, which can arise in a number of forms. First and related to (a), above, polluters may threaten to emit higher levels of pollution in order to secure a larger bribe (Mumey, 1971). Second, as Shoup (1971, pp. 310-312) has pointed out, potential entrants may use extortion: if externality generators are not liable, ‘entrepreneurs’ may threaten to emit an externality in order to secure a bribe, or, symmetrically, if externality generators are liable, potential victims may threaten to ‘come to the harm’ in order to secure a bribe. If these threats necessitate the use of resources to establish credibility, the result will be inefficient. However, investing resources to establish credibility violates the zero transaction cost assumption of the Theorem. Jaffe (1975, p. 661) offers a further counter to Mumey’s result. The polluter will not wish to invest resources in making a threat which will not be carried out and the mere potential that it can carry out a threat will induce a more generous bribe. Moreover, this extortion argument implicitly assumes both sufficient rents and an open class situation and, on the latter ground, is subject to the rebuttal noted above. Demsetz (1972a, p. 23) has also countered this criticism with the argument that competition for these gains will drive the price of extortion to zero, so that extortion is not a barrier to the attainment of the efficient equilibrium. It should also be noted that Coase (1959, p. 27, n. 54) recognized that the employment of resources ‘solely to establish a claim’ could preclude the attainment of the efficient result.

Third, since there are multiple ways to divide the gains from a bargain and each individual is interested in both achieving the benefits from cooperation and getting as large a share of the benefits as possible for himself, there will be threats of noncooperation in order to increase one’s share of the gains. For these threats to be credible, however, they must occasionally be carried out and, when this is done the result will be sub-optimal (Regan, 1972, p. 429). One can see illustrations of what Regan (1972) has called the ‘a priori argument’ for the Theorem in the challenges to the extortion argument. For example, it is argued that the limits of extortion are set by the size of the available rents: if one party tries to extort from the other an amount greater than this, the other party could simply transfer its resources into their next-best use. The extorting agent, being unwilling to forego the potential gain, will thus agree to a solution which
The Coase Theorem

The Coase Theorem garners for it an amount not in excess of the other party’s rents. Thus, the only differential impact of alternative legal rules will be on the distribution of rents; the final allocation of resources will be unaffected (Boyd and Mohring, 1971; Demsetz, 1972a; Feldman, 1974).

(c) Implication Three: Private Information While Davis and Whinston(1965, p. 118) argued early on that information would be revealed through the bargaining process, the application of more complex strategic thinking suggests that private information, if revealed, may be used against one’s self and thus adversely affect one’s payoff. Given this, agents have an incentive to conceal information (through silence or lies) and to expend resources both to protect the value of their own private information and to acquire information from/about others. These costs and resultant delays and/or failures to consummate mutually-beneficial bargains, are likely to preclude the attainment of efficient negotiated solutions where information is asymmetrically distributed (Cooter, 1982; Sutton, 1986; Farrell, 1987). Cooter(1982, pp. 17-18) even goes so far as to argue that an equally strong case can be made that parties will never agree on the distribution of the surplus, even when transaction costs are zero, a proposition that he labels the ‘Hobbes Theorem’. However, he maintains that the ever-present strategic element is not as ‘insurmountable’ as the Hobbes Theorem implies, nor as ‘inconsequential’ as the Coase Theorem implies; in fact, he argues, ‘gains from trade in bargaining situations are realized more often than not’ (Cooter, 1982, p. 19). Cento Veljanovski(1982, p. 60) offers a theorem similar to Cooter’s Hobbes Theorem - the ‘Johansen theorem’, which holds that ‘[d]irect bargaining has an inherent tendency to dissipate the gains-from-trade through strategic behaviour’ (see Johansen, 1979, pp. 515-520). Unlike Cooter, however, Veljanovski maintains that in a world of zero transaction costs the dissipation of gains is likely to be the more common outcome.

A number of commentators have demonstrated the potential for both agreement and non-agreement when information is imperfect. If neither party’s utility function is a function of the other’s private information, then an efficient result will be reached. However, if either party’s utility function is a function of the other’s private information, then there is no guarantee that an efficient result will be reached (Schweizer, 1988, pp. 259-263). (See also Arrow, 1979, pp. 29-31; Cooter, 1982, pp. 20-24; Samuelson, 1985; Illing, 1992). Our discussion here will draw primarily from Cooter’s analysis. Arrow and Samuelson reach conclusions very similar to Cooter, using the assumption that the parties are uncertain about each others’ utility functions. For example, Cooter(1982, pp. 20-24) points out that uncertainty regarding the opponent’s response causes each player to form a rational expectation of this response in the sense of formulating a subjective probability distribution over his opponent’s moves. Given this rational expectation regarding his opponent’s strategy, each player chooses for himself the strategy that maximizes his
expected utility based on a comparison of the greater share of the gains from taking a harder line in bargaining with the higher probability that this harder line will prevent an agreement from being reached. The problem is that, while each player will be playing the strategy that is optimal against the distribution of his opponent’s possible strategies, this strategy is not necessarily optimal against the particular strategy played by the opponent. The outcome will be inefficient when players err in their predictions of the moves made by their opponents (Cooter, 1982, pp. 20, 23; Arrow, 1979, p. 31). In fact, Cooter contends that zero transaction (communication) costs actually decreases the possibility of reaching an agreement, in that it facilitates the transmission of threats and other strategic communications (Cooter, 1982, pp. 23), although, as Arrow (1979) has demonstrated, it is possible to design a collective decision rule that will induce a truthful revelation of preferences.

If the situation involves large bargaining groups, two more potential difficulties arise. First, individuals will have an incentive to free ride and thus the ability of the group to pay a bribe sufficient to induce the socially optimal level of output/pollution will be greatly reduced. Second, if there are differential damage effects across victims, we may observe the rise of coalitions within the victim group (for example, by level of damage), each applying pressure to encourage the result that best suits its interests. The greater is the number of coalitions, the smaller is the likelihood that the optimal solution will be reached (Wellisz, 1964, p. 354). However, as the number of parties approaches infinity (with large numbers of right-holders and large numbers of rights-seekers), the bargaining solution here will approach the efficient result of competitive equilibrium (Samuelson, 1985, p. 338).

In sum, the likelihood of incomplete information gives us little reason to believe that the Coase Theorem is correct when specified in a noncooperative bargaining context. But while the game-theoretic critiques of the Coase Theorem are suggestive of its demise, they have not gone unchallenged, largely on the grounds that it is incorrect to place the Theorem in such a context. At issue is what is meant by a world of zero transaction costs, to which we now turn.

6. The Issue of Transaction Costs

Perhaps the most sticky issue in the debate over the Coase Theorem is the meaning given to the assumption of zero transaction costs. Indeed, the very concept of transaction costs has been so vague and ill-defined that Stanley Fischer (1977, p. 322, n. 5) was once led to remark that ‘almost anything can be rationalized by invoking suitably specified transaction costs’. Coase’s (1960, p. 15) definition of transaction costs encompasses those costs associated with search, negotiation, monitoring and enforcement, which, as Dahlman (1979,
p. 148) has noted, basically reduces to ‘resource losses incurred due to imperfect information’. A bit more specificity (although even greater breadth) is found in more recent definitions within the property rights literature such as Barzel’s (1989, p. 2) contention that transaction costs are the costs associated with the transfer, capture and protection of rights and Allen’s (1991, p. 3) statement that they encompass ‘the resources used to establish and maintain property rights’. Under these latter definitions, zero transaction costs implies complete property rights (Allen, 1991, 1995; Cheung, 1992; see also Schlag, 1989).

If we take the Barzel/Allen definition as the basis upon which to evaluate the Coase Theorem, three implications immediately follow. First, there is an unspecified property right that lies at the heart of the private information examples. For example, Farrell (1987) constructs a problem in which parties A and B care about a time that is set. He notes that King Solomon would have little trouble finding the optimum value since he charges each party an amount equal to their effect on each other and that a ‘bumbling bureaucrat’ can under some circumstances find a superior (second best) result to that achieved under zero cost negotiations. The showing that a centralized authority can achieve an optimum is, however, equivalent to showing that there is an unspecified property right. When A is allowed to set the time (about which both A and B care), B will prefer to not participate in bargaining since his gain is greater when A just sets the time unilaterally and B pays nothing. The advantage that Solomon has is that he can force A and B to participate. Yet, in this example, B gets to use the time that A sets. For example, the time may represent the time for beginning a race. Farrell’s example assumes that B is nevertheless allowed to participate in the race. But, if the race is owned the owner will in fact charge both A and B care, yielding the Solomon solution. Another analogy can be made to ownership of a lake. The owner will charge each user at least the cost they impose on other users. Farrell implicitly assumes that government is the only owner of the lake. It is true that we can consider other examples in which the ownership of the resource seems more foreign to our usual thinking, as when the time represents a curfew or a time after which noise must be reduced—but there is a lack of ownership none the less. These are simply examples in which the absolute advantage of government with respect to certain sorts of enforcement costs may support government ownership, but enforcement costs are also transaction costs.

Second, nearly all of the challenges to the Theorem’s correctness are invalidated under this conception of transaction costs, including entry, nonseparabilities, nonconvexities and even wealth effects, as noted above. This definition of transaction costs also invalidates other challenges to the Theorem which were not discussed in the preceding sections - those based on rent seeking (Jung et al., 1995; see Medema, 1996b) and the presence of risk (Posin, 1990; see Medema, 1995a and Posin’s, 1995, response). Of particular
importance here is the implication of the Barzel/Allen definition (and even Dahlman’s definition) for the game-theoretic challenges, which rely on the existence of imperfect information. By any of these definitions, the presence of imperfect information has the effect of introducing transaction costs into the analysis through the behavior it induces and the game-theoretic challenges are correspondingly invalidated. (On this point see also Allen, 1995, pp. 12-13; Dahlman, 1979, pp. 158-159, n. 26; Hovenkamp, 1990, p. 787; Illing, 1992; Samuelson, 1985, p. 323; Zerbe, 1980, pp. 85-86). Indeed, Saraydar (1983), acknowledges that the imperfect information resulting from strategic behavior violates the assumption of zero transaction costs, but argues that such costs are virtually inevitable within a small numbers bargaining situation due to the incentive to distort information. Further evidence for this line of reasoning may be found within Coase (1960, pp. 31-33), who discusses the problem of moral hazard (along the lines of Harris, 1990) in the context of positive transaction costs.

Against these claims, defenders of the game-theoretic approach point out that giving this type of content to the idea of zero transaction costs basically renders the idea of bargaining meaningless and detaches the Theorem almost completely from reality, making it, in the words of one commentator ‘more in common with astrology than with market analysis’ (Veljanovski, 1982, p. 60). (See also Regan, 1972, pp. 429-430; Shoup, 1971, p. 310; Cooter, 1982, p. 17.) In the end, then, whether the game-theoretic and other challenges to the Coase Theorem go to its correctness or its relevance comes down to ‘how one interprets the almost mystical world of zero transactions costs’ (Zerbe, 1980, p. 85).

This takes us directly to the third implication of the Barzel/Allen definition - that ‘[t]ransaction costs are ubiquitous’ (Allen, 1991, p. 4). The effect of this is to make the Theorem per se completely devoid of applicability to the real world. Coase (1981, p. 187) has made this point a bit more graphically, contending that the analysis of a world of zero transaction costs is akin to ‘divining the future by the minute inspection of the entrails of a goose’. Indeed, by this definition the Theorem’s efficiency proposition must hold, since any violation of it reflects a ‘[cost] associated with the transfer, capture, or protection of property rights’. It may be argued with some justice that all of this reduces the Coase Theorem to a mere tautology (Regan, 1972, pp. 429-30; Cooter, 1989, p. 67). So be it. Coase (1960, p. 15) never claimed that it was realistic, just that it follows logically from the same basic assumptions underlying Pigouvian theory circa 1960. And indeed, based on the foregoing analysis the correctness of the Theorem remains untouched, apart from the potential for taste and preference-induced divergences between WTA and WTP that may impact the invariance claim. The issue of relevance is a different matter altogether and one to which we now turn.
C. Relevance

A further facet of the extensive debate over the Coase Theorem has been the attempt to verify its predictions experimentally and empirically. Hovenkamp (1990, p. 794) has recently pointed out that ‘[c]onducting empirical tests of the Coase theorem is like conducting empirical tests of the Pythagorean theorem. Given the theorem’s assumptions, the results flow out as a matter of logical necessity’. This is true; as such and given the impossibility of satisfying the zero transaction costs assumption in the world in which we live, the experimental and empirical tests go to the issues of relevance and applicability, rather than to the Theorem’s correctness. In doing so, they begin to address the potential difficulties for the Theorem that are raised by the challenges discussed above, such as the effects of imperfect information, the potential for strategic behavior, nonconvexities and the presence of income or taste and preference effects.

7. Experimental Tests

The experimental tests of the Coase Theorem are among the most interesting of the various tests, since they offer the potential to mimic as closely as possible the conditions of zero transaction costs. At the same time, they can begin to capture the effects of factors such as imperfect information and isolate the import of these effects vis-à-vis situations in which they are absent.

7.1 Experimental Framework and Results

The most extensive experimental tests of the Theorem are those undertaken by Hoffman and Spitzer (at times with others - see Hoffman and Spitzer, 1982, 1985, 1986; Coursey, Hoffman and Spitzer, 1987; Harrison et al., 1987). The experiments undertaken by Hoffman and Spitzer all involve the same basic experimental framework. There is a range of possible outcomes, each with a different associated payoff. One party (the ‘controller’) is given the ‘property right’ and thus can determine the outcome unilaterally. Consider the following set of possible dollar payoffs \((P_A, P_B)\) in the spirit of Hoffman and Spitzer: \((5,0)\), \((4,4)\) and \((0,5)\). If A is the controller, he will choose \((5,0)\) unless B induces him to choose a different outcome. The Coase Theorem predicts that \((4,4)\) will be chosen, with the actual distribution of the joint payoff ($8) being a function of the negotiation process. And of course it is in B’s interest to offer A up to $4 to choose (4,4) and in A’s interest to accept any payment greater than $1 to do so.

In well over 500 experiments with various sizes of bargaining groups (including 1-1, 1-3, 2-2, 5-5, 1-9 and 1-19) conducted within this framework, the results were quite favorable. In all, the parties bargained to the efficient result 92 percent of the time, including 94 percent of the time under
conditions of full information and 90 percent of the time under imperfect information. Moreover, the support for the Theorem’s prediction was actually greatest in the 10 and 20 person situations - 98 percent, including 100 percent in 20 person negotiations, where the 19 people devised their own informal institutional arrangements (choosing representatives from the group as bargaining agents) to overcome the large numbers problems. Given how little the student subjects have at stake in these experiments, even those conducted under conditions of imperfect information suggest very low transaction costs. While this may seem quite unrealistic, these experiments do offer some fairly substantial support for the applicability of the Theorem when transaction costs are low.

One of the interesting issues raised by the early experiments came from the nearly equal division of the payoffs (± $1) in the vast majority of the cases. Individual rationality suggests, with respect to the above example, that the controller would not settle for less than $5 and could potentially induce the other party to agree to a $7.99 - $.01 split of the $8 payoff, as in, for example, a one-shot game. The fact that so many controllers settle for less than what should be their reservation price suggests either altruism - which contradicts entirely the theory of externalities, Coase Theorem or otherwise, since, if agents are altruistic, the individually rational behavior that is said to generate the externality in the first place would not occur (Harrison and McKee, 1985, p. 655) - or a potential problem with the experimental environment.

One hypothesis offered to explain this result is that participants in these experiments did not understand the full meaning and import of having a unilateral property right as controllers (Harrison and McKee, 1985). Another is that they did not feel a morally justified right to be the controller, since that position was determined on the basis of a coin flip rather than being, in some (for example, Lockean) sense, earned (Hoffman and Spitzer, 1985). Once measures were implemented to control for this - ‘educating’ subjects or having subjects ‘earn’ the position of controller by winning a preliminary game - the extent of individually rational behavior increased dramatically and without a significant drop-off in the rate at which efficient bargains were made - approximately 90 percent (Harrison and McKee, 1985; Hoffman and Spitzer, 1985, 1986). Even then, however, 20-30 percent of the experiments generated less than individually-rational outcomes, suggesting that the subjects behave more like Lockeans than like utilitarians or egalitarians. Rather than taking this as evidence against the Coase Theorem, Hoffman and Spitzer (1986, pp. 159-160) suggest that it speaks to the robustness of the Theorem across alternative hypotheses regarding individual behavior.
7.2 Criticisms re: Externality Problems

One of the criticisms of these early experiments was that they did not account for the possibility of ‘affronts to dignity’ and other such factors to which people would refuse to assign a monetary value or over which people would refuse to bargain (Kelman, 1985, pp. 1038-1039). In an attempt to deal with the ‘affront to dignity’ issue, Coursey, Hoffman and Spitzer (1987) conducted experiments that introduced a discomforting externality. This was accomplished by introducing the possibility that the ‘victim’ would have to hold one ounce of an unpleasant-tasting liquid in his or her mouth for twenty seconds. These experiments had two possible payoffs: (i) the ‘polluter’ gets $0 and the victim gets $10 and does not have to taste the liquid; (ii) the ‘polluter’ gets $20 and the victim gets $10 and does have to taste the liquid. The latter outcome, where the victim is exposed to the externality, is, of course, the efficient one. Out of 40 experiments, the efficient outcome was chosen 38 times - 22 out of 22 times when the polluter was the controller and 16 out of 18 times when the victim was the controller. The results with the polluter as the controller may not be surprising: when the ‘polluter’ has the right to pollute, polluting is efficient and there are substantial gains to the polluter from polluting, the polluter would be expected to pollute.

What may be a bit surprising, however, is the propensity for victims to sell their right to be free from the harm - here, for a 90 percent increase in payoff (from $10, the payoff to the victim without tasting, to $19.06, the average payoff to the victim when agreeing to taste), or roughly half of the gains to the polluter. In spite of the authors’ claim to the contrary, these results further call into question the assumption of individual rationality (or the inducement thereof within the experimental environment), since, in half of the experiments, the ‘polluters’ paid the victims to taste the liquid even when the polluters had a unilateral right to force them to do so. And while the authors do not jump from a willingness to sell the right to avoid tasting bitter liquid to, say, the legalization of prostitution or pornography (although they do not entirely close the door on such matters, choosing instead to ‘express no opinion’ as to whether such activities should be allowed), they do suggest a presumption in favor of allowing individuals to transfer moderate amounts of dignity and/or moderate amounts of danger.

Based on their various findings, Hoffman and Spitzer (1986, p. 162, emphasis added) assert that their results ‘produce a presumption in favor of the Coase Theorem’, by which they mean that ‘a judge or a legislator should start his analysis by presuming that the parties can and will, in general, exhaust the gains from trade available through private bargaining’ and that those who would argue against this ‘must bear the burden of the proof’. Furthermore, they argue, the strength of the evidence for the optimality of the bargaining outcomes establishes a presumption in favor of injunctive over damages
remedies because of the high probability that the parties will bargain to the efficient result (Hoffman and Spitzer, 1986, pp. 163-168). (See Calabresi and Melamed, 1972. For discussions of the relative efficacy of property rules and liability rules when transaction costs are positive, see also Polinsky, 1979; 1980; Ayres and Talley, 1995a, 1995b; Kaplow and Shavell, 1995, 1996.)

In making these claims, Hoffman and Spitzer clearly go too far. It is one thing to show that the Coase Theorem is largely confirmed in a laboratory setting that attempts more or less to mimic the zero transaction costs world, but it is quite another to say that the results thus generated establish ‘a presumption in favor of the Coase Theorem’ for efficiently resolving real-world externality problems with twenty or fewer parties. (It should be noted that the parties will always reach an efficient point, in the Pareto sense. That is, even if transaction costs are so large as to preclude rights transfers, that result is Pareto efficient, given transaction costs. See, for example, Samuels (1974), Buchanan (1983) and Calabresi (1991). However, the efficient result to which Hoffman and Spitzer’s ‘presumption’ refers is the Pigouvian social optimum, which is equivalent to the Pareto optimal outcome when transaction costs are zero.) What may be established is that up to twenty-party externalities will be resolved efficiently through negotiations in many and perhaps even most, instances where transaction costs are very low and the stakes are very small. But transaction costs consist of far more than factors introduced by a twenty-second tasting of a foul liquid or adding parties to a bargain, particularly when the group consists of more-or-less homogeneous college students. The emphysema-ridden residents of the neighborhood are likely to have a far different view of pollution externalities than would others who are not so affected; the light sleepers are likely to look far differently at the neighborhood kennel than are the deep sleepers and so on. And how many groups are absent one or two members who are likely to impede any negotiated settlement? All of this is to say nothing of the information and coordination problems that may attend complicated real-world bargains. In sum, Hoffman and Spitzer develop some very nice results offering rather strong support for the Theorem when the conditions it assumes are nearly met. However, to move from this to claims of widespread applicability and presumptions in favor of the Theorem in real-world cases is a somewhat different matter, one requiring far more caution and future study than is implied by the authors.

7.3 Criticisms re: Invariance Proposition
Apart from transaction-cost-related issues of applicability, these experiments also fail to get at the invariance proposition per se. For example, might effects such as the normative sanction for rights, an affront to dignity, or wealth effects cause a divergence between WTA and WTP? These issues are side-stepped within the Hoffman and Spitzer experimental design that makes efficiency and
invariance go hand-in-hand. While the evidence regarding the existence of a divergence between WTA and WTP is not uniform (see the survey by Hoffman and Spitzer, 1993, and the references cited therein), it seems to weigh in favor of the existence of a non-trivial divergence, one that is too large to be explained solely in terms of wealth effects. Attempts to measure the relationship between WTA and WTP have been done through surveys and through experiments. The survey evidence shows a substantial difference between WTA and WTP, although economists have been rather suspicious of these results because of their several potential biases. However, as we have noted, recent experimental treatments of this issue support the contention that WTA may be substantially greater than WTP, often more than twice as great (Levy and Friedman, 1994; Hoffman and Spitzer, 1993, pp. 69-85). While this literature is too vast to survey here, one set of experiments, undertaken by Kahneman, Knetsch and Thaler (1990), probes this issue in the context of the Coase Theorem and, in a number of different types of experiments, finds significant endowment effects. Using items such as mugs, pens, binoculars and chocolate bars, they find that individuals, when given the opportunity to exchange these items for cash, exhibit a strong reluctance to part with entitlements and thus that, contrary to standard assumption of economic theory, preferences are apparently not independent of entitlements (Kahneman, Knetsch and Thaler, 1990, p. 1339). The value that the subjects place on these objects ‘appears to increase substantially as soon as the individual is given the object’ (Kahneman, Knetsch and Thaler, 1990, p. 1342) and the resulting disparity between WTA and WTP does not dissipate in repeated trials (that is, with market experience). They suggest that this endowment effect is most likely to occur for items that are not easily replaceable, which makes the endowment effect particularly important for the Coase Theorem, since things like a nice view, or clean air or water, are not easily replaced and it thus can be expected that people will refuse to sell such goods even at prices somewhat greater than their reservation price for buying them.

Two implications of these WTA versus WTP experiments are particularly important. The most obvious is that the results provide strong evidence against invariance in the outcomes of bargains even when transaction costs are zero. Second, endowment effects reduce the gains from trade as compared with a world in which preferences are independent of endowments. Since fewer mutually advantageous exchanges are possible, the volume of trade is lower than it otherwise would be (Kahneman, Knetsch and Thaler, 1990, p. 1344). Given the size of the potential disparity between WTA and WTP, one can conceive of situations where each party’s WTA is greater than the other party’s WTP, so that no trade would occur, whereas if WTA were equal to WTP, we would see bargains consummated. Experiments run to test this implication in a Coase Theorem context revealed substantial under-trading relative to the Theorem’s predictions (Kahneman, Knetsch and Thaler, 1990, pp. 1339-1341).
Even when transaction costs are negligible, then, there does not seem to be much room for confidence in the generation of an invariant outcome when consumers are party to an externality.

### 7.4 Criticisms re: Different Contexts

Even less favorable to the Theorem’s applicability are the results of an experiment by Stewart Schwab (1988), who looked at the Theorem in the context of labor law and labor-management negotiations. The graduate student subjects in Schwab’s experiments were asked to negotiate a union contract over wages, vacation time and - the crucial aspect of the experiment - whether or not the company had the right to transfer work to its nonunion plant over the course of the three-year contract. The implications for the Coase Theorem lay in the contract presumption that was said to govern labor relations in the absence of a specific contract provision: in one group of experiments, subjects were told that the legal presumption was that the company must continue to use union workers unless the contract explicitly states otherwise (that is, includes a ‘go clause’), while the other group was told that the presumption was that the company could transfer work to the nonunion plant during the course of the contract unless the contract explicitly stated otherwise (that is, includes a ‘stay clause’). (These contract presumptions have actual counterparts in labor law B the *Milwaukee Spring* cases.)

Analysis of the results of these experiments shows that only about 20 percent of the contracts were fully efficient when wage levels, vacation time and the stay or go clause are accounted for, a vast difference from the roughly 90 percent efficiency of the Hoffman and Spitzer and the Harrison and McKee experiments. Out of 108 contracts, all but two had efficient wage levels, but only 31 percent had efficient vacation levels and only about 65 percent had a stay clause where it was efficient or a go clause where it was efficient. Schwab contends that three factors may account for these differences. First, parties in these experiments were bargaining over multiple contractual terms under a binding time constraint and thus may have found it difficult to make efficient choices on all items. Second, unlike many of the other experiments, subjects here were not given full and perfect information. This meant both that information had to be communicated during the negotiation process and that signaling and bluffing could occur, leading to inefficient agreements. Finally, the subjects did not know what the ‘best’ outcome was from the beginning and thus had to find their way to it and do so over a rather large bargaining range which, of course, could easily result in inefficient outcomes (Schwab, 1988, pp. 251-252).

Given these factors, the rather high rate of failure to reach efficient bargains is not particularly surprising. The environment of these experiments corresponds much more closely to a natural setting than do many of the other experimental treatments and the factors that exist in these natural settings are
such as to weigh heavily against the attainment of efficient results. By the same token, however, these experiments go much more to the application of the insights of the Coase Theorem than to the testing of the Theorem \textit{per se}, in that they introduce a variety of factors that are assumed away by the Theorem’s assumption of zero transaction costs. Thus, earlier work, such as that by Hoffman and Spitzer, speaks favorably to the Coase Theorem on its own terms, while Schwab’s results are rather pessimistic about the ability of parties to bargain to efficient results in more natural settings. However, Schwab’s results take on a better cast when evaluated in light of what he calls the ‘weak efficiency hypothesis’, which says that the law will not affect the \textit{rate} at which efficient bargains are consummated. (This is not to be confused with the ‘weak’ version of the Coase Theorem, noted above, which asserts efficiency but not invariance.) The form of the contract presumption does not have a significant effect on the inclusion of stay clauses or go clauses, nor on the efficiency of the contracts (Schwab, 1988, pp. 252-253).

8. Empirical Studies

There have been three studies that more or less take Coase’s farmer-rancher example into the real world to look at the ability of parties to negotiate efficient solutions to animal trespass problems.

8.1 California Animal Trespass Laws 1850-90

Kenneth Vogel (1987) examines the response of farmers and ranchers to changes in California animal trespass laws between 1850 and 1890. At the time when, in 1850, California joined the Union, its principal industries were mining and cattle raising and, reflecting the importance of the cattle industry to the state, California had what was, in essence, strict nonliability for cattle trespass. This rule clearly favored the ranchers and the evidence strongly suggests that the rule was designed with that in mind. At the same time, however, it played a major role in hindering the development of agriculture in the state (Vogel, 1987, pp. 163, 167).

However, between 1851 and 1890 there were no less than 150 different laws enacted by the California legislature altering the rules that governed cattle trespass in ways that benefited farmers (Vogel, 1987, pp. 163-164). The Coase Theorem predicts that these alterations in the law will have no effect on the allocation of resources; that is, \textit{ceteris paribus}, these changes in the law should have no effect on the level of resources devoted to ranching and farming, or on ranching and farming outputs. And, according to Vogel, this situation is particularly well-suited to testing the applicability of the Theorem to the real world since ‘the externality is visible, the parties are, at least post hoc, easily
identifiable and it is easy to measure, or use proxies to estimate, the damages’ (Vogel, 1987, p. 181).

Contrary to the Theorem’s prediction, however, the enactment of the various estray laws beginning in the 1860s was accompanied by an enormous increase in farm output, particularly for wheat farming, which became common in the valleys, while cattle were moved up into the foothills. Econometric analysis undertaken by Vogel shows that a number of variables attempting to capture the effects of legal change on crop outputs are significant and that these are uniformly positive in sign, which supports the claim that these legal changes did indeed influence the growth of agriculture (Vogel, 1987, p. 184).

Given the strength of the evidence, it remains to explain why the Coase Theorem fails here. Vogel suggests two reasons. First, transaction costs may be significant and, furthermore, are asymmetric across alternative assignments of rights; specifically, they are lower when ranchers are liable (Vogel, 1987, pp. 176, 187). When ranchers are not liable for the damage done by their cattle, the farmer wishing to keep the cattle off his land will have to negotiate with each rancher whose cattle might potentially stray onto his land in order to accomplish this. On the other hand, when farmers have the right to compensation for damages, the onus is on the ranchers to initiate negotiations and the rancher need only bargain with those farmers on whose land his cattle may be expected to stray. Second, there are nonconvexities as a result of the externality. If ranchers are not liable for trespass damages, the fact that each farmer has to negotiate with all ranchers whose cattle might stray onto his land in order to prevent damage means that ‘[i]f a farmer fails to contract with the rancher whose cattle actually use his land, all payments made to the other ranchers are useless’ (Vogel, 1987, p. 176). The farmer will thus have little incentive to initiate such negotiations, with the result that efficiency will obtain only if it is efficient for cattle to be allowed to roam freely. This nonconvexity is not present when farmers are given the right (Vogel, 1987, pp. 174-176, 187). Taken together, these two factors can explain why output was lower when the ranchers were not liable.

In contrast to Hoffman and Spitzer, who use their experimental results to claim a presumption in favor of the Coase Theorem, Vogel (1987, pp. 186-187) argues that his results refute the general applicability of the Theorem. Yet, Vogel has not refuted the Theorem but rather has shown the importance for the case of straying cattle of transactions costs and pointed out the importance of assigning the legal rule to minimize transactions costs as both Coase (1960, p. 19) and Posner (1983, p. 71) suggest should be done. Given the enormous volume of legal change at the time, the nonconvexities present and the difficulty of ascertaining the source of the damage, Vogel’s broad conclusion may be premature. Some degree of support for a more moderate view can be found in the study of contemporary trespass disputes undertaken by Robert Ellickson.
8.2 Effects of Open- v. Closed-Range Laws
Ellickson (1986, 1991) examines, among other things, the effects of open-versus closed-range laws on cattle trespass disputes in Shasta County, California. Under open-range laws, cattlemen are not usually responsible for accidental trespass damage, whereas they are strictly liable under closed range laws. Ellickson finds that cattlemen and their neighbors do in fact behave in a manner suggested by the Coase Theorem, cooperating to resolve their disputes regardless of who is liable. However, the evidence also suggests that it is not Coase Theorem-type mechanisms at work here; rather, individuals seem to rely on community norms to determine their behavior. For example, while the Theorem predicts that the cattleman would install a fence if he were liable (closed range) and that the neighboring farmer would do so if he were liable (open range), it is almost always the cattleman who installs the fence because both cattlemen and their neighbors believe that the cattleman is morally obligated to do so, since his cattle cause the damage. Moreover, the citizens seem to be very ignorant of the relevant law and ignore those aspects of the law that conflict with their view of the world. As such, they do not bargain ‘in the shadow of the law’ (see Mnookin and Kornhauser, 1979, and Cooter, Marks and Mnookin, 1982), but beyond it; community norms seem to have much more force than the legal rule in place. Ellickson suggests that this may be due to the fact that relations among the neighbors are both complex and continuing, because of which the transaction costs associated with acquiring information and litigating disputes are high and reliance on norms offers a lower-cost way of resolving these disputes. Ellickson (1989, 1991) also suggests that this norm-based behavior points to the need to revise certain of the behavioral concepts underlying law and economics.

8.3 Roaming Deer in Scottish Highlands
In a study that has interesting commonalities with that of Ellickson, Nick Hanley and Charles Sumner (1995) examine an externality situation owing to the roaming of red deer in the Scottish Highlands which cause damage to growing trees and, in the process, impose substantial costs on the owners of these forests, the value of the timber from which is diminished. In addition, the wandering deer may destroy growing crops on farmland and, when they stray onto sheep grazing land, reduce the forage for sheep, thus imposing costs on both farmers and sheep ranchers. The beneficiaries of the red deer population are estate owners, who derive substantial income and estate value from the presence of red deer on their estates (Hanley and Sumner, 1995, pp. 88-91).

Given the level of damage, the small number of parties, the ease of quantifying damage to forests and the relative ease with which estate owners could reduce the size of their herds, the situation seems to reflect an inefficiently-high deer population and a fertile ground for the working of Coase
The Coase Theorem-type mechanisms. Even so, an extensive study by Sumner (1993) failed to turn up any instances of Coasean bargaining between owners of deer estates and neighboring landowners. What one does observe, however, are Deer Management Groups which neighboring landowners have established ‘to coordinate deer management across neighboring estates ... and forest/farmland, in order to reduce the level of the externality’. The advantage of such groups is that they ‘effectively [internalize] the externality across members of the group’, thereby avoiding the third-party effects that can result with bilateral bargaining (Hanley and Sumner, 1995, p. 93). It is interesting to note the parallel between the rise of the cooperative Deer Management Groups and the behavior of neighbors revealed in Ellickson’s study of cattle ranching in Shasta County. While the law offers a low-cost option (free government culling) for dealing with red deer damage, groups of neighboring landowners in essence ignore the law and work out a solution amongst themselves, perhaps presumably because it is the case that the transaction costs associated with the cooperative efforts of the Deer Management Groups are lower than those that would attend bilateral negotiations of the Coasean variety (Hanley and Sumner, 1995, p. 93).

8.4 Implications for other Legal Rules: Divorce

The Coase Theorem has implications for all manner of legal rules, including, as Peters (1986, 1992) points out, the rules governing divorce. Since 1970, there has been a progressive movement in the US from divorce by mutual consent (requiring the agreement of both parties), to unilateral divorce, where the marriage can be terminated at the demand of either party. Intuitively, the rules governing divorce function to establish property rights with respect to dissolution of the marriage. Under unilateral divorce law, the spouse seeking divorce has property rights with respect to dissolution while, under mutual consent, the right rests with the spouse who does not wish to see a divorce occur (Peters, 1992, p. 690). The conventional wisdom was that unilateral divorce laws would make divorce easier (in economic terms, reduce transaction costs), thus increasing the divorce rates in states that adopted such laws. The Coase Theorem predicts that if bargaining costs are minimal and information is symmetric across parties, divorces will only be undertaken when they are efficient (that is, joint benefits exceed joint costs), regardless of the law governing divorce and that the legal rule will have no impact on the divorce rate.

Peters (1986) tests a model corresponding to the Coase Theorem environment against one that posits asymmetric information and thus predicts that divorce rates will differ across alternative legal rules and finds that the data support the predictions of the Theorem against the conventional wisdom: the move to a unilateral divorce rule does not affect the probability that a woman becomes divorced (Peters, 1986, pp. 446-448). Moreover, the level of alimony and child-support payments are ‘significantly lower’ in states with unilateral
divorce rules (Peters, 1986, p. 449) and the labor force participation rates of married women in such states are higher, which, she argues, may represent an attempt by married women to self-insure against the possibility of becoming divorced without compensation (Peters, 1986, pp. 448-449, 451-452). Thus, both the divorce rates and the distribution of compensation are consistent with the Coase Theorem. However, a number of subsequent studies find that the evidence tends to support the conclusion that divorce rates are in fact higher in unilateral divorce states than in mutual consent states. (See, for example, Allen, 1992; Zelder, 1993a, 1993b; Brinig and Buckley, 1995; Friedberg, 1995, which also contain numerous citations to literature on both sides of the argument.) Within this debate, each side claims that the other’s empirical work contains errors or biases that influence the results (see, for example, Peters, 1986, 1992; Allen, 1992).

Even if one is willing to accept the result that divorce rates are not impacted by the legal rules governing divorce, it remains to ascertain whether these results actually reflect the working of Coase Theorem-type mechanisms. There is plenty of reason to suggest that the answer is ‘no’, or at least ‘not necessarily’. First, there is no evidence to suggest that ‘unilateral’ divorces are undertaken only when they are efficient. Rather, Peters infers that the efficiency proposition holds based on a questionable claim that transaction costs are low and the fact that the data confirm the model’s invariance and distribution predictions. (For contrasting views on the potential magnitude of transaction costs here, see Peters (1992, p. 690) and Allen (1992, p. 684). Allen (1992, 1995) goes so far as to argue that, by working an uncompensated transfer of wealth from wives to husbands, the move to no-fault divorce violates the zero transaction costs/fully-specified rights condition assumed by the Theorem and thus that the rise in the divorce rate does not constitute a legitimate argument against the Theorem.) Furthermore, Peters fails to account for the fact that the Theorem predicts not just an invariant divorce rate, but an invariant allocation of household resources as well - just as the farmer-rancher example predicts an invariant allocation of resources devoted to farming and ranching. The fact that female labor force participation is higher in states with unilateral divorce rules thus speaks loudly against the claim of invariance.

8.5 Implications for Other Legal Rules: Pre-Trial Settlements
A similar problem attends the claim that the high rate at which suits are settled prior to trial supports the Coase Theorem (Hoffman and Spitzer, 1986, pp. 168-169). Glanter (1983, pp. 28-30) finds that roughly 90 percent of all lawsuits are settled before they go to trial and that, when they do not settle, it tends to be due to (i) cases that require a judicial decree to be settled; (ii) cases that are not costly to litigate, which reduces the incentive to settle; (iii) the placing by one or more parties of special value on having a judicial decree for
reasons including, _inter alia_, precedent and reputation; (iv) cases that involve an issue that is not easily negotiated over, such as a ‘fundamental value’; and (v) the high transaction costs associated with settlement as compared to going to trial - factors that lie outside of the bounds of the Coase Theorem. However, the high proportion of settlements does not imply that the parties have bargained to the Pareto optimal result contemplated by the Coase Theorem. There is no way to infer from the settlement data whether the parties have bargained to the socially optimal result contemplated within a zero transaction costs world, or if they simply have realized some of the potential gains from negotiation but hit a point where the transaction costs from further negotiation exceed the expected gains and choose to settle at a ‘suboptimal’ outcome because this settlement, although not optimal, is still better than going to trial.

8.6 Implications for Other Legal Rules: Unemployment

Perhaps the most unique empirical test of the Coase Theorem is Donohue’s (1989a) analysis of the Illinois employment experiment (Spiegelman and Woodbury, 1987), which attempted to determine whether the payment of bonuses to unemployed workers for securing employment, or to employers for hiring unemployed workers, would reduce the duration of unemployment and the costs associated with the unemployment compensation system. While the experiment was conducted to ascertain how such bonuses might affect the duration of unemployment, its application to the Coase Theorem is straightforward. The efficiency hypothesis predicts that mutually advantageous bargains will be struck under either scheme and all workers and employers who satisfy the eligibility requirements will collect bonuses. The allocative invariance proposition suggests that members of the worker-payment group (WPG) will find jobs and collect bonuses at the same rate as members of the employer-payment group (EPG). The invariant distribution hypothesis predicts that members of the WPG and EPG groups will have the same aggregate compensation (wages plus bonus). The WPG workers would be expected to have lower wages than EPG workers, reflecting a bargaining away of a share of their bonus, as compared to EPG workers’ higher wages as employers bargained away a share of their bonus.

Donohue’s inquiry into the results of the Illinois experiment reveals that they contradict the predictions of the Coase Theorem on all counts. To begin with, the number of bonuses paid to WPG workers was about five times that paid to EPG employers. Furthermore, many workers and employers who met the requirements for bonuses failed to submit a voucher to receive their bonus - particularly employers. Given this, says Donohue, ‘[t]he conclusion that a number of individuals and employers acted inefficiently is hard to rebut’ (Donohue, 1989a, p. 573). The experimental results also revealed that members of the WPG had a significantly greater improvement in obtaining employment, relative to the control group, than did members of the EPG, which, along with
the differential bonus collection rates among the groups violates the allocative invariance prediction of the Theorem (Donohue, 1989a, pp. 569-577). Finally, there was no significant difference in wages across the WPG and EPG hires, so that aggregate (wages plus bonus) compensation was higher for WPG workers than for EPG workers, in violation of the invariant distribution prediction of the Theorem. That is, it appears that workers did not bargain with employers over the wage or the bonus as a result of this program (Donohue, 1989a, pp. 586-590).

Donohue (1989a, pp. 591-601) asserts that transaction costs are extremely low here and thus that the Theorem fails in a case most favorable to its success. He offers two possible explanations for why the experiment failed to satisfy the predictions of the Coase Theorem, each of which goes to the issue of individual behavior/decision making in the context of such bargains. One possibility, to which Donohue lends a great deal of support, is ignorance on the part of workers, who seemed not even to realize that bargaining was possible and may not have understood that the bonuses could aid them in gaining employment. Moreover, there may have been an authoritarian effect which caused the workers to believe that the bonus was in effect inalienable - that the party designated to receive the bonus was in fact entitled to its full value, an effect that would also discourage the type of bargains envisioned by the Theorem (Donohue, 1989a, pp. 600-602).

Donohue’s conclusions regarding the implications of the Illinois experiment for the Coase Theorem’s applicability have been challenged by Ellickson (1989) and Lindgren (1990), both of whom contend that Donohue greatly underestimates the effect of transaction costs within this experiment. Acquiring this so-called ‘free’ money actually involves a substantial number of steps (and even more for employers than for workers), the aggregate effect of which is to make the process rather costly, relative to the size of the bonus. Lindgren (1990, pp. 581-582, 585), for example, lists the steps that participants in the WPG and EPG programs must go through in order to collect bonuses and offers several reasons why one would not anticipate an invariant distribution of income, or bargaining over the bonuses. In fact, Lindgren (1990, p. 583) suggests that the results do in fact match the predictions of the Coase Theorem (see also Ellickson, 1989, p. 625).

First, because costs are high relative to the bonuses, the insights underlying the Coase Theorem would lead one to predict that many workers or employers who are eligible for bonuses would not collect them. The Illinois study confirmed this prediction. Second, because costs are higher for employers than for workers, the insights of the Coase Theorem would lead one to predict that more workers in the worker-bonus group would be influenced to participate in the program, obtain work quickly and collect the bonuses. Again, the Illinois study confirmed these predictions (Lindgren, 1990, p. 583).
Indeed, if transaction costs were zero, the entire premise of the experiment would disappear, since job search is a positive transaction cost phenomenon (Lindgren, 1990, p. 578). Finally, the presence of stigma and institutional rigidity effects would lead one to predict minimal bargaining over wages and bonuses (in an amount different from the control group), a result again supported by the Illinois experiment. In fact, Ellickson finds the results of the Illinois experiment very consistent with his own study of Shasta County - that people often tend to rely on norms rather than ‘legal rules’ to govern their behavior, particularly when the stakes are low and there is an expectation of a continuing relationship, as among neighbors in Shasta County and between employers and employees in the Illinois experiment (Ellickson, 1989, pp. 627-628).

8.7 Implications for Other Legal Rules: Tenancy

The two earliest attempts to empirically validate the working of Coase Theorem-type mechanisms in real-world environments were undertaken by Cheung (1969a, 1973), who examined two of the classic illustrations of market failure caused by externalities - share tenancy arrangements in agriculture and the relationship between beekeepers and apple orchard owners. These studies undertake to examine the ‘Stigleresque’ version of the Theorem, which asserts the efficient internalization of externalities under conditions of perfect competition.

The standard view in the economics and tenancy literatures has long been that share tenancy leads to an inefficient allocation of resources, owing to (i) the short duration of the leases; (ii) the discouragement of effort on the part of the tenant, since a portion of each unit of output must be paid to the landowner as rent; and (iii) the disincentive for either party to make investments in the land that will maximize the land’s productivity (see Cheung, 1969a, pp. 3-4, 7-8 and the references cited therein). As a result, one would expect to observe lower crop yields under share tenancy than under alternative cultivation arrangements. However, the Coase Theorem predicts that, if transaction costs are zero and there are well-defined and freely alienable private property rights in land, the allocation of resources will be ‘the same whether the landowner cultivates the land himself, hires farm hands to do the tilling, leases his holdings on a fixed rent basis, or shares the actual yield with his tenant. In other words, different [observed] contractual arrangements do not imply different efficiencies of resource use Y’ (Cheung, 1969a, p. 4).

An examination of share tenancy in China and Taiwan prior to the land reforms of 1949 shows that there was a well-developed system of private property rights in land in China and Taiwan at this time and that the market by and large comported with the dictates of competition. Furthermore, Cheung did not observe lower ratios of labor and other inputs, a lesser degree of improvements, or lower yields on tenant farms than on owner-cultivated farms.
or on farms employing wage labor, nor is there evidence that the market values of land under tenant cultivation are lower than the values of land under owner cultivation (Cheung, 1969a, pp. 56-61, 1980, p. 42). And, while it has been argued that, under share tenancy, certain types of activities - such as improvements to farms - would be contracted inadequately or not at all, Cheung (1980, p. 43) finds that these ‘are precisely the activities stated in every written contract that I could find’. All of this evidence suggests quite strongly that charges that share tenancy is less efficient than other cultivation arrangements cannot easily be sustained.

Even though the theory assumes zero transaction costs where transaction costs are actually positive, it is able to explain ‘much of the observed farming behavior’ (Cheung, 1969a, p. 159). In this situation of well-defined property rights, transaction costs are not so high as to affect resource use at the margin, but, rather, affect the choice among alternative contractual arrangements - the use of alternative methods of cultivation reflecting the tradeoff between coordination costs and risk, each of which varies across alternative contractual arrangements (Cheung, 1980, p. 44). (For a related discussion, see Cheung, 1969b.) While Cheung’s evidence does not conclusively demonstrate the optimality of share contracts, it certainly does lend strong support for the claim that, at a minimum, share contracting is no less efficient than other available contractual arrangements and thus that, under the appropriate (and not unrealistic) conditions, the mechanisms of the Coase Theorem can lead to a satisfactory resolution of externality problems through the market.

The other early study by Cheung of the working of Coase Theorem-type mechanisms is ‘The Fable of the Bees …’ (1973), a study which responds to Meade’s (1952) classic discussion of the positive reciprocal externalities that exist between beekeepers and the owners of apple orchards: apple blossoms provide valuable services to beekeepers, whose bees feed on them, while, at the same time, bees provide valuable pollination services to the apple-orchard owner. While Meade argued that a system of taxes and subsidies can and must, be imposed in order to achieve efficiency, contractual arrangements between farmers and beekeepers have long been routine in the US and the existence of a market for nectar and for pollination services can be readily observed in the state of Washington, the location of Cheung’s study - in some cases merely by consulting the yellow pages of the telephone directory (Cheung, 1973, p. 19). The question, of course, is whether these markets generate an efficient allocation of resources. Cheung (1973, pp. 24-28) argued that a presumption can be established in the affirmative, since available data provides substantial support for the competitive nature of the market.

How is it that in this externality situation the market avoids the failure pointed to by Meade? To begin with, transaction costs are very low here. Since the value of resources devoted to pollination and nectar extraction is
insignificant and farmers could easily and cheaply keep bees themselves (and sometimes do so), the gains from contracting with beekeepers are extremely small, which, in turn, suggests that contracting costs are minimal (Cheung, 1980, pp. 46-48). There is also a well-developed system of contractual relations between beekeepers and farmers, so well-developed, in fact, that, while written contracts (sometimes as simple as postcards) are used to secure an initial arrangement among the parties, oral agreements are standard for subsequent relations. Furthermore, these oral contracts are rarely breached, owing to the presence of ‘extra-legal constraints’ in the form of sanctions against those who do not honor their contracts (Cheung, 1973, p. 29). Yet, in spite of the informality of these contracts, they tend to be quite comprehensive, specifying ‘the number and strength of the [bee] colonies, the rental fee per hive, the terms of delivery and removal of hives, the protection of bees from pesticide sprays and the strategic placing of hives’. And, where hives are placed merely for honey-generating purposes (that is, no pollination is involved), prices (often paid in honey) are not necessarily fixed - being allowed to vary with the honey yield (Cheung, 1973, p. 29). All of these various pieces of evidence lead Cheung to conclude that, contrary to Meade’s story, ‘the allocation of hives and nectar flows approximates that of a smoothly functioning market’ wherein resources are efficiently allocated (Cheung, 1980, p. 50).

This having been said, Cheung notes that there are two factors which could potentially complicate these arrangements (relative to standard lease contracts), both of which relate to other levels of external effects. First, there are potential spillovers from one farmer contracting for pollination services, which could potentially lead neighbors to take strategic advantage by employing fewer hives themselves. Second, the use of pesticide sprays by one farmer may result in damage to the bees kept on nearby farms. But both of these issues are dealt with through either custom or explicit contracting (such as the payment of risk premiums for potential exposure to pesticides), depending on the circumstances. The reliance on customs here is an interesting parallel to Ellickson (1986, 1991), discussed above.

It should be obvious that it is not possible to confirm or refute the efficiency claims made by Cheung and, given this, the results cannot be said to show the applicability of the Theorem (here, the Stigler, 1966, version) per se. Yet, they offer important evidence that markets can successfully (if not fully efficiently) deal with potential externality problems under the appropriate conditions.
D. Importance

9. The Importance of the Coase Theorem

The importance of the Theorem lies not in whether it is correct, but in the detailed nature of the assumptions required to make it correct-in that is the nature of transactions costs. The interesting question about the Theorem then is the nature of the transactions costs associated with those situations in which the theorem is thought to not be correct. The legacy of the Theorem lies in the subsequent work attempting to detail the nature of transaction costs and their effect on the workings of the economic system. The richness and variety of types of economic arrangements can now be seen in the richness and variety of transaction costs and in mechanisms for reducing them.

In this sense the importance of the Coase Theorem lies not in its supposed correctness or incorrectness and the corresponding policy relevance or lack thereof, but, rather, in the positive transactions cost propositions that flow from it. These include both normative and positive propositions. Examples are Posner’s (1992, p. 52) normative suggestion that at law ‘[s]ince transactions are never costless in the real world, efficiency is promoted by assigning the legal right to the party who would buy it ... if it were assigned initially to the other party’ and the positive prediction of Lesser, Dobbs and Zerbe (1997) that suits at law in situations where negotiation costs are low will involve considerations of distribution not efficiency.

The Coase Theorem has helped to give rise to an extensive body of work, much of it summarized by Eggertsson (1990), concerned with economic behavior and institutions and to a more detailed and useful sense of what is meant by property. The Coase Theorem has made clearer the relationship between transactions cost and property rights and in doing so has begun to give a much stronger basis for understanding how legal regimes change in response to changes in constraints (North, 1981). One can now define the strength of property rights in terms of lower transaction costs for the exclusion, exchange and use of property.

In part of the economics literature at least (Eggertsson, 1990) the transactions cost approach has replaced the market failure model of public intervention that is expressed by Weimer and Vining (1992, p. 30):

When is it legitimate for government to intervene in private affairs? In the United States, the normative answer to this question has usually been based on the concept of market failure - a circumstance where the pursuit of private interest does not lead to an efficient use of society’s resources or a fair distribution of society’s goods.
When the costs of transaction mechanisms are introduced there are departures from perfect markets. These departures are in fact externalities because they represent effects not taken into account in the decision making process. Externalities, then, are found everywhere there are transaction costs and are ubiquitous. Since the concept of market failure rests on externalities that are defined by transactions costs, the concept of market failure (and the concept of externality) does no work for us that is not already done by transaction costs (Zerbe and McCurdy, 1996). Externalities are in fact an unnecessary complication in the theory of government intervention.

Public goods represent a useful example of a situation in which the market failure model can be and to some extent is being, replaced by a transaction cost model. The older market failure approach is represented by Samuelson (1954), who saw public goods as a class of market failures. For example, in writing about the classic case of the lighthouse, Samuelson (1964, p. 45) writes:

Here is a later example of government service: lighthouses. These save lives and cargoes; but lighthouse keepers cannot reach out to collect fees from skippers. So, says the advanced treatise, ‘we have a divergence between private advantage and money cost ... and true social advantage and cost ... Philosophers and statesmen have always recognized the necessary role of government in such cases of ‘external-economy divergence between private and social advantage’.

The transaction cost-property rights approach appears to provide a richer vehicle of analysis, as shown by Cheung, North and others. For example, Coase (1974a) shows that the British lighthouse system was once a well-functioning private system and that, in general, the system was more complex than that suggested by the simplistic market failure diagnostic.

The triumph of the transaction cost approach shows that the true legacy of the Coase Theorem lies not in its correctness, but in drawing attention to the role played by transaction costs within the economic system.

10. Conclusion

In light of the foregoing discussion, three things can be said about the Coase Theorem. First, it is correct, in the sense that it has withstood all of the challenges mounted against it to date. Efficiency will obtain, regardless of the initial assignment of rights and the result will be invariant keeping in mind that even income effects are irrelevant if one accepts the Barzel/Allen definition of zero transaction costs. Second, the Theorem, although correct, is unrealistic, as Coase recognized (see Coase 1960, p. 15, 1981, p. 187, 1988a, pp. 174-179). The latter point, of course, should have been obvious from the beginning, which
The Coase Theorem raises the question as to why the debate over the Theorem has been so intense. A small part of the answer, within the economics profession at least, may lie in the interesting theoretical puzzle that the Theorem poses. Dwarfing this, however, is the normative debate that, implicitly or explicitly, pervades nearly all aspects of the Theorem’s discussion.

Three prescriptions for legal-economic policy are said to flow from the Coase Theorem.

1. Rights-cum-market solutions are said to be preferable to Pigouvian remedies for the resolution of externality problems.
2. Property and contract are efficient; any interference with the outcomes so generated will make matters worse rather than better. It is this implication that makes the Coase Theorem, in the minds of some, ‘the cornerstone of a laissez-faire legal and economic policy regarding contract and property law’ (Hoffman and Spitzer, 1986, p. 151).
3. When transaction costs are positive, rights should be assigned to those who would possess them in the end-state if transaction costs were zero, as seen in the prescriptions of wealth maximization, or ‘mimic the market’.

But the Coase Theorem says none of these things. The Theorem is a positive statement with no normative implications; it is an ‘is’ statement, not an ‘ought’ statement. Each of the above propositions rests on the assumption that efficiency is the goal of legal-economic policy. But the Coase Theorem goes merely to the presence of absence of efficiency; it does not tell us that it is all that matters or even that it matters at all. It is this normative leap that seems to underlie most of the hostility to the Coase Theorem - and, by extension, to law and economics generally. (See, for example, Baker, 1975; Kelman, 1979; the Symposium on Efficiency as a Legal Concern, 1980; A Response to the Efficiency Symposium, 1980; Schlag, 1986; Gjerdingen, 1986; Johnston, 1990 and Crespi, 1991.) For a response to these types of criticisms see Zerbe (1998a).

Compounding the hostility to the normative use of the Coase Theorem are the use of incorrect definitions of economic efficiency by proponents of law and economics (Zerbe, 1998b).

Furthermore, even if one takes efficiency to be the goal of legal-economic policy, the Coase Theorem does nothing to establish the sanctity of property and contract or the superiority of the market over Pigouvian remedies, owing to the ubiquitous nature of transaction costs. If coordination is costless, the market will optimally allocate rights and resources, but so too will Pigouvian remedies. On the other hand, the Pigouvians fare no better in this debate since, after waving away the Coase Theorem on the grounds that transaction costs are positive, they tend to immediately fall back on the demonstration that Pigouvian remedies generate socially optimal outcomes, using models in which
government is assumed to operate with full information and without cost. (But see Baumol, 1972, for a more judicious evaluation of Pigouvian remedies.)

Here, we come to the true import of the Coase Theorem. The Theorem is not, in the end, about markets or about costless bargaining; rather, it is about the costs of coordination. If coordination is costless, markets function perfectly; but so does government. If coordination is costly, markets function imperfectly; but so does government. The task for legal-economic policy thus becomes that of ascertaining the magnitude and influence of these costs and the resulting implications for alternative institutional-policy arrangements. The true and valuable legacy of the Theorem is all of the subsequent work on transactions costs that explore the costs of coordination under different regimes and in different situations. Like the Coase Theorem itself, this, too, is without direct normative implications. However, it has led to certain normative claims, as noted above. One, from an analytical perspective, is that the received conception of externalities should be abandoned. Another, this time from a policy perspective, is that judgments as to the appropriate form of government intervention should be made on the basis of what institutional arrangement produces the lowest combination of coordination costs. In this regard, it is interesting to note that much of the normative debate and the propositions we noted above in that regard can be turned into a series of positive predictions about which arrangements will promote economic efficiency. Where the affected parties could reach a solution through negotiation but choose litigation or regulation, the real issue is likely to be who is to be assigned property rights rather than how to realize gains from trade (see, for example, Lesser, Dobbs and Zerbe, 1997).

Finally, the meaningfulness of the Coase Theorem must be understood in epistemological terms. The ‘correctness’ of the Theorem is a matter of logical validity; in general, the Theorem is a conclusion derived from premises and the role of the assumptions constituting its premises is to rule out of consideration all those variables which would prevent the derivation of the conclusion as a matter of logic. The validity of the Theorem, therefore, is a function of the assumptions defining away certain limiting conditions. The empirical truth of the Theorem - its descriptive accuracy - is a separate matter from its logical validity. The Theorem considered empirically is a tendency statement, a statement that under certain conditions such and such behavior and allocative and so on results can be expected; that is, a law in the Marshallian sense. However, the logical and empirical aspects are closely related to one another in that changing the assumptional conditions of the Theorem is tantamount to changing the conditions in terms of which the Theorem is a tendency statement. The Theorem is a tendency or probability statement in a further empirical sense, to wit: the experimental literature indicates that the results expected on the basis of certain specifications of the Theorem are realized something less than one hundred percent of the time.
Given the foregoing, it becomes clear that much of the literature on the Coase Theorem not only overreaches in terms of the normative implications more or less improperly drawn from the Theorem, but fails to specify the meaningfulness of the Theorem in such terms - readily leading to practices which make claims for and take the Theorem far beyond what logicality and empiricism permit.

Acknowledgments

The authors would like to thank Douglas Allen, James Buchanan, Robert Cooter, Robert Ellickson, Elizabeth Hoffman, Richard Posner, Warren Samuels and participants in the Workshop in the History of Economic Thought and Methodology at Michigan State University for comments and suggestions on previous drafts of this chapter and for helping to clarify our thinking on certain points raised herein. The excellent research assistance of Mary Therese Cogeos is also gratefully acknowledged. Of course, the authors are solely responsible for any remaining errors.

Bibliography on The Coase Theorem (0730)


The Coase Theorem


Glanter, Marc (1983), 'Reading the Landscape of Disputes: What we Know and Don’t Know (and Think we Don’t Know) about Our Allegedly Contentious and Litigious Society', 31 UCLA Law Review, 4-71.


More, Daniel (197), ‘The Coase Theorem and the Reciprocal Nature of Land-Use Conflicts in Pollution Cases’, 4 Studies in Law, Tel Aviv University, 84-114.


Ng, Yew Kwang (1975), ‘Coase’s Theorem and First-party Priority Rule: Reply’, 51 Economic Record, 272-274.


The Coase Theorem


Schweizer, Urs (1988), ‘Externalities and the Coase Theorem: Hypothesis or Result?’, 144 Journal of Institutional and Theoretical Economics, 245-266.  

Cases

UAW v. NLRB, 765 F.2d 175 (D.C. Cir. 1985).