



ELSEVIER

Economics Letters 61 (1998) 3–11

**economics  
letters**

# The Coase theorem is tautological, incoherent or wrong

Dan Usher\*

*Queen's University, Department of Economics, Kingston, K7L 3N6, Canada*

Received 2 February 1998; accepted 16 April 1998

---

## Abstract

The Coase theorem is commonly understood to mean that costless bargaining ensures efficiency in the economy for any assignment of property rights. The standard demonstration of the theorem suggests that costless bargaining ensures efficiency without an assignment of property rights. © 1998 Elsevier Science S.A. All rights reserved.

*Keywords:* Property; Pricing; Bargaining and Coase

*JEL classification:* B41

---

'If transaction costs are zero, the initial assignment of a property right – for example, whether to the polluter or to the victim of pollution – will not affect the efficiency with which resources are allocated.' The Coase theorem, as stated by Richard Posner

'The Problem of Social Cost' (Coase, 1960) contains no explicit Coase theorem, but commentators have professed to see such a theorem, the only theorem to my knowledge with an established name but no universally-recognized content. In 'Coase Theorem' (Eatwell et al.), Robert Cooter argues that the theorem is 'false or a tautology' but for somewhat different reasons than those put forward in this paper. So far as I can tell, Cooter interprets bargaining as the act of negotiating and he argues that, though negotiation itself may be costless, no bargain need be achieved. I, on the other hand, interpret bargaining as the successful striking of a deal, in which case negotiation without agreement entails a transaction cost in the wastage of the surplus over which the negotiation takes place. On the definition in this paper, failure to agree is itself a transaction cost. I have no quarrel with Cooter's arguments. My purpose in this paper is: (i) to show that the Coase theorem is a very questionable entity, and (ii) to clarify the inextricable connection between price-taking and bargaining, one determinate within economic analysis, the other not.

A widely-accepted statement of the Coase theorem is that of Richard Posner above (Posner, 1993). Thus stated, the Coase theorem is doubly ambiguous. The words 'will not affect' are open to two distinct meanings with entirely different implications about property rights, and the term 'transaction cost' conceals a fundamental assumption about bargaining. These ambiguities will be considered in turn.

The form of the theorem is 'If A, then B', where the antecedent is

\*Tel.: +1 613 5452285; fax: +1 613 5456668; e-mail: usher@qed.econ.queensu.ca

A: ‘that transactions cost are zero’,

and the consequent is

B: ‘that the initial assignment of a property right will not affect the efficiency with which resources are allocated’.

The statement B is open to two interpretations. It may be interpreted as

$B_1$ : ‘that *any* assignment of property rights gives rise to an efficient allocation of resources, on the understanding that efficiency requires *some* assignment of property rights’,

or it may be interpreted as

$B_2$ : ‘that resources will be allocated efficiently regardless of whether or not there is an assignment of property rights’.

The distinction between  $B_1$  and  $B_2$  is of fundamental importance in economics and in the law where the implications of the Coase theorem are supposed to lie. The statements  $B_1$  and  $B_2$  are mutually exclusive; they cannot both be correct. Interpreted as ‘If A, then  $B_1$ ’, the Coase Theorem implies that, in the absence of transaction cost, a clear delineation of property rights is a necessary and (subject to well-known qualifications not relevant here) sufficient condition for efficiency in the economy, though any apportionment of property to people will do as well as any other. Interpreted as ‘If A, then  $B_2$ ’, the Coase theorem implies that, in the absence of transaction cost, property rights are unnecessary because people bargain their way to efficiency in the economy regardless. In the latter interpretation, bargaining becomes a perfect substitute for property and the price mechanism.

Clearly, it is the proposition ‘If A, then  $B_1$ ’ that is enunciated in Posner’s version of the Coase theorem. It turns out, however, that the proposition ‘if A, then  $B_1$ ’ is, strictly speaking, false, while the proposition ‘If A, then  $B_2$ ’ is true but grossly misleading and almost tautological. To establish these claims, I turn to Coase’s famous example of the farmer and the cowboy. Coase constructs this example in terms of marginal revenue and marginal cost. While not wrong, the construction is inexpedient and perhaps misleading. The example is reconstructed here with the aid of a production possibility frontier from which the appropriate marginal curves could, but will not be, derived.

Recall the example. The cowboy’s ranch is situated beside the farmer’s farm. If unattended, cows from the ranch graze on the farm, destroying crops. The destruction of crops is costly. It is also costly to restrain the cows. Costs of crop damage and of restraining cows are both functions of the number of cows that wander onto the farmer’s field. Incomes of both parties are determined accordingly. The income of the farmer is maximized when there are no cows in his field. The income of the cowboy is maximized when there are four cows in the farmer’s field. The value of production – the sum of the income of the cowboy and the income of the farmer – is maximized when there are two cows in the farmer’s field.

The incomes of the farmer and the cowboy are illustrated on Fig. 1, with the income of the farmer, F, on the vertical axis and the income of the cowboy, C, on the horizontal axis. Each of the five numbered points on the figure shows the earnings of the farmer and the cowboy as determined by the number of cows that graze on the farmer’s land. The points labeled 0, 1, 2, 3, and 4 show the incomes

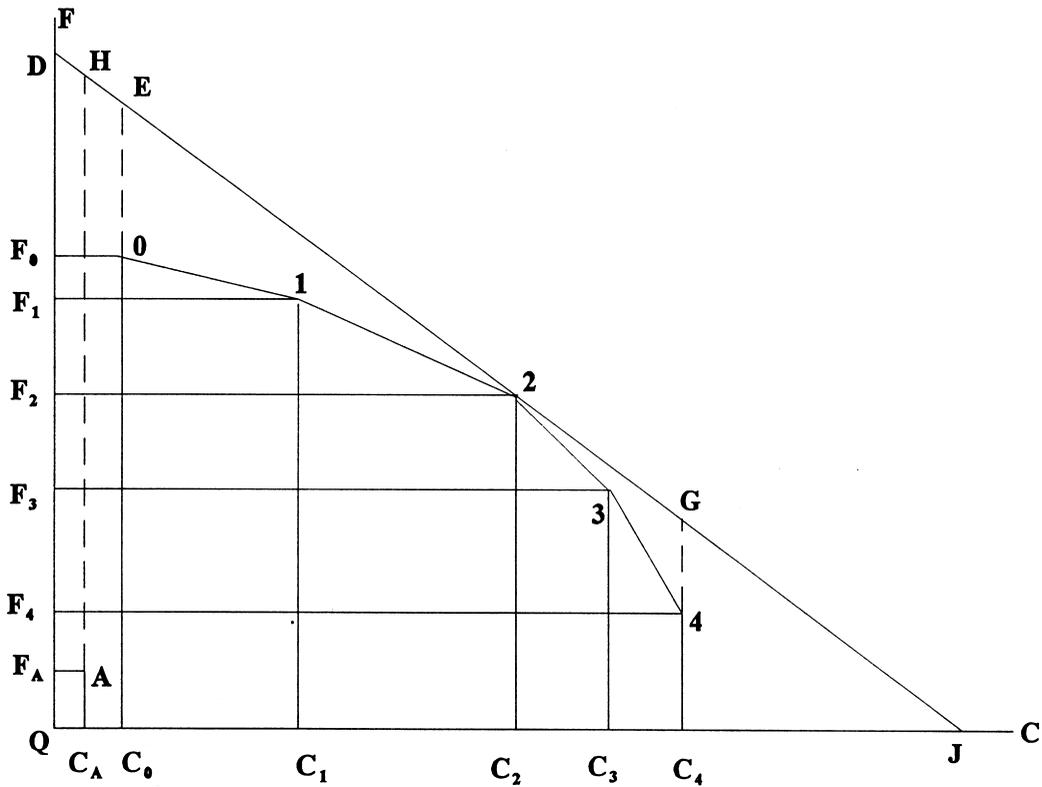


Fig. 1. Income of the farmer ( $F$ ) and income of the cowboy ( $C$ ) depending on the number of cows in the farmer's field.

of the farmer and the cowboy when there are no cows in the farmer's field, one cow, two cows, three cows and four cows respectively. With no cows in the farmer's land, the farmer's income is  $F_0$  as indicated on the vertical axis, and the cowboy's income is  $C_0$  as indicated on the horizontal axis. With some cows in the farmer's field, the incomes of the farmer and the cowboy are indicated accordingly. It would not be inconsistent with the diagram if farming were no longer profitable with as many as four cows on the farmer's field; in that case  $F_4$  would be the farmer's best alternative after giving up farming altogether.

Together, the set of points 01234 is the production possibility frontier for the farmer and the cowboy when outputs are measured in dollars' worth. Their combined income corresponding to any point on the production possibility frontier is the intersection with the vertical axis or the horizontal axis of a line through that point at  $45^\circ$  to both axes. The efficient point on the production possibility frontier is that for which the combined income is as large as possible. It is immediately evident from the figure that the efficient point is 2 with two cows in the farmer's field and incomes to farmer and cowboy of  $F_2$  and  $C_2$ . The line  $DJ$  – at  $45^\circ$  to both axes – is the locus of all possible apportionments between the farmer and the cowboy of their total income at the efficient number of cows in the farmer's field. Necessarily, the maximal combined income is

$$QD = QJ = F_2 + C_2 \tag{1}$$

The gap between ‘the actual combined income of the farmer and the cowboy for any number of cows permitted to graze on the farmer’s field’ and the ‘maximal potential combined income’ (when two cows graze there) can be measured as the vertical distance below the line DJ of the indicator of the incomes of farmer and cowboy with the initial allocation of property rights and in the event that the cowboy is not persuaded or bribed to graze fewer cows on the farmer’s land than the law allows. The potential gain from cooperation is E0 when the initial allocation is at 0, and is G4 when the initial allocation is at 4.

Coase makes the valid argument that, with costless bargaining, the efficient output is attained no matter what the initial assignment of property rights. Suppose the farmer has the absolute right to stop cows from grazing on his land. Without the farmer’s permission, no cows may graze there. By forbidding grazing altogether the farmer can earn  $F_0$  and the corresponding income of the cowboy is  $C_0$ . But it would be disadvantageous for the farmer to exercise that right completely. Instead, the farmer sells the cowboy the right to graze two cows on the farmer’s land, raising their combined income from  $F_0 + C_0$  to  $F_2 + C_2$ , and generating a surplus  $S_0$

$$S_0(F_2 + C_2) - (F_0 + C_0) = E0 \quad (2)$$

which is divided between them in accordance with the price they set. Farmer and cowboy both become better off as long as the price paid by the cowboy to the farmer is greater than  $F_0 - F_2$ , at which the entire surplus accrues to the cowboy, and less than  $C_2 - C_0$ , at which the entire surplus accrues to the farmer. Farmer and cowboy divide the surplus by choosing a payment from among the set of feasible payments for which both parties are better off than if there were no deal at all. To say that bargaining is costless is to say that there is no loss of resources or time when a price within the feasible range is chosen.

Their post-bargain incomes are  $F_\alpha$  and  $C_\alpha$  where

$$F_\alpha = F_0 + \alpha S_0 \quad \text{and} \quad C_\alpha = C_0 + (1 - \alpha) S_0 \quad (3)$$

and where  $\alpha$  is the mutually-agreed upon farmer’s share of the surplus from the bargain. Necessarily,  $0 < \alpha < 1$ . The farmer gets the entire surplus if  $\alpha = 1$ . The cowboy gets the entire surplus if  $\alpha = 0$ . Farmer and cowboy are presumed to agree on some  $\alpha$  between these limits, but nothing in economic theory enables one to predict where within these limits the agreed-upon  $\alpha$  will be. The initial allocation of rights sets bounds on the final allocation of income, but does not determine it uniquely. The set of possible incomes of the farmer in all of the mutually-advantageous bargains is represented by points on the line OE above the allocation in the initial assignment of property rights. One cannot say a priori which point within that range will be agreed upon.

Similarly, if the cowboy has the property right to graze as many cows as he pleases on the farmer’s land, he could, by exercising that right to the full, earn an income of  $C_4$ , leaving the farmer with an income of only  $F_4$ . Once again, it is in the interest of the party with the property right in grazing to sell part of that right because it is worth more to the other party than it is to him. The cowboy sells the farmer the right to have no more than two cows grazing on his land. The surplus is now  $S_4$  where, as is evident from the figure,

$$S_4 = (F_2 + C_2) - (F_4 + C_4) = G4 \quad (4)$$

As the figure is drawn,  $S_4 < S_0$  indicating that the surplus from the bargain is less than it was before.

Now the payment – this time from the farmer to the cowboy – must lie within the range from  $F_2$ – $F_4$ , at which the entire surplus accrues to the cowboy, to  $C_4$ – $C_2$ , at which the entire surplus accrues to the farmer. One way or another and regardless of the initial allocation of rights, the farmer and the cowboy agree to produce efficiently and to allocate of the maximal combined income as represented by some point on the line DJ.

The moral Coase draws from the story is that, with costless bargaining, the efficient output is attained regardless of who has the initial assignment of rights. The total value of production is maximized in either case. Only the distribution of income is affected by the initial assignment of rights. The cowboy gets the larger share of total income when he has the initial property right to graze cows on the farmer's land. The farmer gets the larger share of total income when he has the initial property right to forbid grazing on his land. In either case, there is a surplus to be allocated by bargaining.

What appears not to have been recognized is that the argument proves too much. The argument is a demonstration of the proposition 'if A, then B', as A and B are defined at the outset of this paper. Interpreting the statement B in the sense of  $B_1$ , the proposition might be looked upon as a lesson about how property rights promote efficiency in the economy. But the proposition 'If A, then  $B_1$ ' turns out to be completely false. Property rights are unnecessary in this context. Accept the premise of costless bargaining, and an efficient output can be attained not just for any initial allocation of property rights, but without property rights at all!

What looks at first as a demonstration of the proposition 'If A, then  $B_1$ ' is in reality a demonstration of the proposition 'If A, then  $B_2$ '. To appreciate the irrelevance of property rights when bargaining is altogether costless, imagine what would happen if the cowboy and farmer find themselves side by side with no property rights assigned. Do they fight? Perhaps. Regardless, something must happen, and whatever that something is or whatever the uncertainty about the eventual outcome, there must for both parties be a distribution of possible outcomes with certainty equivalent incomes represented, for the cowboy, by the point  $C_A$  in the figure and, for the farmer, by the point  $F_A$ . Together, these certainty equivalent incomes are represented on the figure by the point A, which is mnemonic for anarchy.

To speak of the expected incomes of the farmer and the cowboy in the absence of property rights is not to deny that the lives of the farmer and the cowboy in these conditions would be solitary, poor nasty, brutish and short. The point A is placed close to the origin to suggest precisely that possibility. But, however badly off the farmer and the cowboy may be in conditions of anarchy, there must be some pair of incomes  $F_A$  and  $C_A$ , such that the parties would be as well off with those incomes held securely as they would expect to be in conditions of anarchy. Think of those incomes as the certainty equivalents of expected incomes net of the cost of conflict between farmer and cowboy when property rights are insecure and it is not known in advance which party will prevail.

Once again, the assumption that farmer and cowboy can bargain costlessly implies that two cows will be allowed to graze in the farmer's field, yielding an efficient outcome at the point 2 where their incomes would be  $F_2$  and  $C_2$  if there were no side payments. The surplus in the bargain becomes  $S_A$  where

$$S_A = (F_2 + C_2) - (F_A + C_A) = AH \quad (5)$$

which must somehow be allocated between farmer and cowboy if the efficient outcome is to be

attained. With costless bargaining, an allocation of the surplus is agreed upon and the value of production is maximized. The difference between this case and the preceding cases is one of magnitude not of kind.

The proof that efficiency is attained in the absence of property rights is qualitatively identical to the proof that efficiency is attained regardless of how property rights are assigned. If one of these propositions are valid, then the other must be valid too. Furthermore, though the proof of this assertion is constructed for two parties only, the logic of the proof remains valid when the number of parties is increased without limit. In a wide variety of circumstances, costless bargaining renders property, and the price mechanism, unnecessary.

Returning to the interpretation of the Coase theorem, the example demonstrates the proposition 'If A, then B<sub>1</sub>' to be false and the proposition 'If A, then B<sub>2</sub>' to be true. With costless bargaining, there is no need for property rights. But, if the proposition 'If A, then B<sub>2</sub>' is what the Coase theorem is supposed to mean, then the theorem as stated is grossly, almost ludicrously, misleading. Recall the wording of B in the original statement of the theorem: 'that the initial assignment of a property right will not affect the efficiency with which resources are allocated'. Reading that, one would not infer that property rights are irrelevant for efficiency. If 'If A, then B<sub>2</sub>' is what is meant, the original version of the Coase theorem is like the statement, 'In the present state of medical technology, painting of fire engines blue rather than red will not change the fact that all men are mortal'. Strictly speaking, that is true, but the statement carries the implication that the painting of fire engines and the mortality of mankind are somehow connected, that, perhaps, men would become immortal if fire engines were not painted at all. As enunciated by Posner, the Coase theorem is incoherent in that it suggests one thing and means another which is almost the opposite.

The other ambiguity in the statement of the Coase theorem – that the term 'transaction cost' is ambiguous – can now be discussed briefly. The word 'cost' bears the implication that one can acquire something for a price. A car may cost \$30 000. A building may cost \$30 000 000. To speak of cost normally implies a well-defined number of dollars, or of some commodity, that must be given up to acquire something else. Transaction cost is not like that. The term is used in the context of bargaining problems where a surplus is to be allocated between parties if and only if they can agree upon shares. Suppose the farmer has the right to exclude all cows from his land and, as indicated in Fig. 1, there is a surplus of GE to be shared between the farmer and the cowboy. A transaction cost would be like an ordinary cost if there were a well-specified mechanism causing total output to fall by some given proportion of GE as a consequence of the bargaining procedure. That is precisely what is absent in this context. Resources may well be used up in bargaining, but there is no telling in advance what the value of those resources will be. Men may kill one another over trifles, or may allocate huge amounts by a simple offer and acceptance. There is in practice no simple one-to-one relation between the magnitude of the surplus over which people bargain and the resources used up in deal-making. There is no rational procedure, comparable the price mechanism in competitive markets, for allocating surplus between interested parties. Surpluses do get allocated, but the process by which they are allocated is, for the economist, fundamentally mysterious. There are plenty of solutions to the bargaining problem, but, to the best of my knowledge, every solution is obtained by adding structure to the crude bargaining problem as set out in the text. For instance, the Nash Bargaining solution is obtained by maximizing the product of utilities, and the Rubinstein bargaining solution is obtained by imposing a sequence of offers and rejections together with a postulated shrinking of the pie over time. See Osborne and Rubinstein.

Talk of transactions cost is, in my opinion, the adoption or imposition of a rule of thumb which is not unreasonable in itself but should be recognized as such. The rule of thumb is that a cost of bargaining exists and is a well-defined function of the surplus to be divided. The rule is that the cost of bargaining,  $B$ , is an increasing and convex function,  $f$ , of the surplus  $S$ ,

$$B = f(S) \quad (6)$$

where  $f' > 0$  and  $f'' > 0$ . To complete the story, there is usually added an assumption about how the surplus net of transaction cost is apportioned between the parties to the bargain. For example, if the initial allocation of property rights is that the farmer may deny all grazing of cows on his land, if the cost of bargaining is  $f(S) = (0.1)S^2$  where  $S$  is the surplus from the bargain, if  $S = OE$  in Fig. 1 and if the residual surplus is divided equally between farmer and cowboy, then the bargain will be such that the farmer's income becomes  $F_0 + OE[1 - (0.1)(OE)^2]/2$  and the cowboy's income becomes  $C_0 + OE[1 - (0.1)(OE)^2]/2$ , as long as  $OE[1 - (0.1)(OE)^2]/2$  is positive. To repeat: There is no denying that bargaining problems are frequently resolved in practice. To look upon bargaining as a well-defined transaction cost is a useful rule of thumb. The rule is ungrounded in any general specification of rational self-seeking behaviour.

Some concluding observations:

(1) That resources will be employed efficiently in the absence of transaction cost is almost a tautology. What does it mean to say that transaction cost are zero? It means that people can bargain costlessly, and will, presumably, do so as long as bargains are mutually advantageous. But the absence of mutually-advantageous bargains is precisely what one means by efficiency, a state of affairs such that no change can make one person better off without making somebody else worse off. The strictly-correct version of the Coase theorem boils down to the proposition if people can agree upon an efficient outcome, then there will be an efficient outcome.

(2) I have no quarrel with Ellickson's defense of Coase (Ellickson, 1991) that 'The essence of Coase's argument...is that transaction costs are large and that economic actors arrange their institutions with an eye to these costs.' My only reservations are about the ambiguity in the concept of transaction cost and that there is no room within that defense for anything that might reasonably be called the Coase theorem.

(3) A distinction can be drawn between price-taking and deal-making. One way or another, resources must be moved from those who have them to those who can make the best use of them. Savings must somehow be transformed into investment. Tomatoes must pass from the farmer who grows them to the people who wish to eat them. The skill of the computer programmer must be placed at the disposal of the bank seeking to automate its services. The price mechanism is the ideal moving van if and in so far as people are price-takers, that is, if products are standardized (as one of a number of well-specified commodities) and everyone acts as though he believed he could buy or sell any amount of a commodity without affecting its price. The core propositions in the theory of general equilibrium are that (i) price-taking behaviour generates an equilibrium where everything for sale is purchased, and (ii) the equilibrium is efficient in the sense that no reallocation of goods or factors of productions could make anybody better off without making somebody else worse off. Universal price-taking is a reasonable assumption for an economy with well-defined goods and many traders in every market, so that nobody, all by himself, can affect any market significantly.

But price-taking is not universal. The market may not discover the equilibrium prices. Goods may

not be standardized. Businesses are unique, and participants with different skills must cooperate in circumstances where market prices cannot be identified for all transactions. Neighbours' interests are directly opposed over the range of permissible uses of property. All too frequently, actors in the market confront one another in actual (as distinct from what economists call perfect) competition, where one's income depends on how shrewdly he conducts himself in face-to-face negotiation with other people.

In most formulations, the theory of general equilibrium is about how prices guide production and distribution efficiently once property rights are established. By contrast, a world with costless bargaining has no need of either property or prices. An essential difference between price-taking and deal-making should be recognized. Both, if they work as they should, yield outcomes that are Pareto optimal, outcomes such that no change can make some people better without making others worse off. The difference is that, with minor exceptions which do not concern us here, the outcome of price taking is unique, while the outcome of deal-making is not. For any given assignment of resources to people, price-taking leads to one and only one outcome, while deal-making leads to one out of an infinite number of possible Pareto optimal outcomes, with no basis for predicting which outcome that will be. The point is well-known, but is worth restating here.

(4) The two-person model in this paper tends to conceal an important aspect of property rights. Property facilitates bargaining by restricting the number of participants to the bargain. An implicit assumption in this paper, as in Coase's original telling of the story of the farmer and cowboy, is that bargaining is between the farmer and cowboy alone. Nobody else gets to participate. With no property rights, the butcher, the baker and the candlestick-maker would all demand a share not just of the surplus, but of the total income from farming and cattle raising together; and, with costless bargaining, their demands could somehow be accommodated with no loss of output at all. If bargaining were really and truly costless, the entire population of the world would cooperate to maximize world income which would then somehow be shared amicably, in a great *reductio ad absurdum* of the major premise of the Coase theorem. It is because bargaining is never costless and because the difficulties of bargaining increase with the number of parties involved that economies must rely primarily upon pricing and upon command. The world's work gets done through prices, with bargaining at the edges where property rights conflict or unique resources have to be combined in a common enterprise.

(5) Both versions of the Coase theorem – 'If A, then B<sub>1</sub>' and 'If A, then B<sub>2</sub>' – require that agreements, once reached, will be respected, either because the parties to the agreement are trustworthy or because agreements are enforced by the state. In the latter case, it would be odd if a government prepared to enforce private contracts was not also prepared to establish property rights, but the contrary is imaginable if not realistic. Settlers occupy a territory without prior property rights, and the government stands ready to enforce whatever rights to property and whatever contracts the settlers establish among themselves.

Interpreted as a general reminder that the economy runs on a mixture of price-taking and deal-making and that bargaining is no free good, the Coase theorem is instructive but misnamed as a theorem. Interpreted to mean that costless bargaining promotes efficiency in the economy, the Coase theorem is a tautology, for a bargain among rational people must make each person better off than he was before and, with costless bargaining, self-interested people bargain again and again until no mutually-advantageous bargain remains. Interpreted as implying that efficiency in the economy requires an allocation of property to people, even when bargaining is costless, the Coase theorem is incoherent or wrong.

## **References**

- Coase, R.A., 1960. The problem of social cost. *Journal of Law and Economics*, 1–44.
- Eatwell, J., Milgate M., Newman, P. (Eds.), *The New Palgrave: A Dictionary of Economics*, vol 1. pp. 457–60.
- Ellickson, R.C., 1991. The case for Coase and against Coaseanism. *Yale Law Journal*, 99, 611–33, p. 612.
- Osborne, M., Rubinstein, A. *Bargaining and Markets*.
- Posner, R., 1993. Nobel laureate: Ronald Coase and methodology. *Journal of Economic Perspectives*, Fall 1993, 195.