6-1-1999

How Coasean Bargaining Entails a Prisoners' Dilemma

Wayne Eastman

Follow this and additional works at: http://scholarship.law.nd.edu/ndlr

Part of the Law Commons

ARTICLES

How Coasean Bargaining Entails a Prisoners' Dilemma

*Wayne Eastman**

As typically presented, the Coase Theorem and the Prisoners' Dilemma are very different stories with very different messages.[1] The Coase Theorem, whether expressed in Coase's terms of a cattle rancher and a farmer[2] or in another way,[3] is associated with a conservative, Chicago School message which stresses the ability of rational, self-interested parties to reach an optimal outcome through bargaining, without the aid of regulation or liability standards.[4] The Prisoners' Dilemma, on the other hand, whether expressed in its original version involving two prisoners or in another version,[5] is associated with a liberal message which emphasizes the

---

1 For discussions of how economic models such as the Theorem and the Dilemma combine logic, metaphor, and storytelling, see DONALD N. MCCLOSKEY, IF YOU'RE SO SMART: THE NARRATIVE OF ECONOMIC EXPERTISE (1990); DONALD N. MCCLOSKEY, THE RHETORIC OF ECONOMICS (1985); Donald N. McCloskey, *The Rhetoric of Law and Economics*, 86 MICH. L. REV. 752 (1988). For a discussion of the standard stories of the Coase Theorem, the Prisoners' Dilemma, and supply-demand equilibrium, along with a presentation of heterodox stories, see Wayne Eastman, *Telling Alternative Stories: Heterodox Accounts of the Prisoners' Dilemma, the Coase Theorem, and Supply-Demand Equilibrium*, 29 CONN. L. REV. (forthcoming 1996).

2 Ronald H. Coase, *The Problem of Social Cost*, 3 J.L. & ECON. 1, 2-8 (1960).

3 *See, e.g.,* RICHARD A. POSNER, ECONOMIC ANALYSIS OF LAW 8 (1992) (illustrating Coase's theorem using the example of a railroad emitting sparks onto adjacent land); GEORGE J. STIGLER, THE THEORY OF PRICE 113 (3d ed. 1966) (mentioning formalized statement of theorem).

4 For discussions of the politics of the Coase Theorem, see C. Edwin Baker, *The Ideology of the Economic Analysis of Law*, 5 J. PHIL. & PUB. AFF. 3 (1975); Guido Calabresi, *Transaction Costs, Resource Allocation and Liability Rules—A Comment*, 11 J.L. & ECON. 67 (1968) (arguing for a liberal interpretation of the Theorem); Donald H. Gjerdingen, *The Politics of the Coase Theorem and Its Relationship to Modern Legal Thought*, 35 BUFF. L. REV. 871 (1986); Herbert Hovenkamp, *Marginal Utility and the Coase Theorem*, 75 CORNELL L. REV. 783 (1990); Mark Kelman, *Consumption Theory, Production Theory, and Ideology in the Coase Theorem*, 52 S. CAL. L. REV. 669 (1980); Pierre Schlag, *An Appreciative Comment on Coase's* The Problem of Social Cost: *A View from the Left*, 1986 WIS. L. REV. 919 (interpreting the Theorem as potentially a vehicle for radical analysis of legal concepts). Although they see the Coase Theorem as a viable vehicle for liberal or radical politics, Calabresi, Gjerdingen, and Schlag acknowledge the basic point made in the text that the Theorem's background is in free market, anti-interventionist "Chicago" politics.

5 For discussions of the Dilemma, see ROBERT AXELROD, THE EVOLUTION OF COOPERATION (1984) (discussing repeated Prisoners' Dilemmas); DOUGLAS R. HOFSTADTER, METAMAGICAL THEMAS 715-55 (1985) (discussing Axelrod's results and advocating a cooperative approach to the single-round Dilemma); R. DUNCAN LUCE & HOWARD RAIFFA, GAMES AND DECISIONS 94-102 (1957) (presenting formally-oriented treatment of the Dilemma and other games); PARADOXES OF RATIONALITY AND COOPERATION (Richmond Campbell & Lanning Sowden eds., 1985) (collecting articles by philosophers); ANATOL RAPOPORT, TWO PERSON GAME THEORY (1966) (including reflections on the meaning of "rationality"); Martin Shubik, *Game Theory, Behavior, and the Paradox of the Prisoner's Dilemma: Three Solutions*, 14 J. CONFLICT RESOL. 181 (1970).

inability of self-interested parties unaided by legal or moral controls to reach an optimal outcome.[6]

Given the dissimilarity of the Coase Theorem and Prisoners' Dilemma stories—and the sharp difference between their received messages—one might well assume that the Theorem and the Dilemma would also be dissimilar when expressed formally. But one would be wrong: Coasean bargaining, expressed in the form of a 2 × 2 matrix, constitutes a Prisoners' Dilemma matrix.[7]

That is the formal statement of my point.[8] Given the widespread diffusion of both the Coase Theorem and the Prisoners' Dilemma stories in legal academia[9] and more generally in contemporary American intellectual culture,[10] a demonstration of their formal kinship is a project whose value need not be argued.[11] I note, however, that the formal point is also of interest because it is related to other issues. Specifically, in thinking about law (and other subjects), one faces a choice between relying on stories such as the Prisoners' Dilemma and the Coase Theorem, which are told as illustrations of certain logical principles, or relying on stories that employ logic in a less structured fashion and assign it a less prominent place than do the

---

6 *See* HOFSTADTER, *supra* note 5; RAPOPORT, *supra* note 5. Both Hofstadter and Rapoport offer personal statements concerning the liberal politics of the Dilemma.

7 By putting the proposition this way, the point is not to claim a priority for the Dilemma over the Theorem. Because the Dilemma has from its early days in the 1950s been expressed formally as a 2 × 2 matrix, and the Theorem has not, the Dilemma is in the position of being there first for purposes of analysis in matrix form. In addition, as I will show, the matrix expressing Coasean bargaining is a special case of the Prisoners' Dilemma matrix; all Coasean bargaining can be expressed in the form of a Prisoners' Dilemma matrix, but the reverse does not hold. *See infra* note 33. For these reasons, the proposition in the text (and this Article's title) relates the Theorem to the Dilemma, rather than vice versa.

8 To the best of my knowledge, the point that Coasean bargaining constitutes a Prisoners' Dilemma is one that has not been made in Coase Theorem or Prisoners' Dilemma literature. In formulating this point, literature on the formal relationship of the Prisoners' Dilemma to Newcomb's Problem (a puzzle involving a superbeing with predictive power) has been helpful to me. *See generally* David Lewis, *Prisoners' Dilemma Is a Newcomb Problem, in* PARADOXES OF RATIONALITY AND COOPERATION, *supra* note 5, at 251; J. Howard Sobel, *Not Every Prisoner's Dilemma Is a Newcomb Problem, in* at 263.

9 LEXIS searches found 509 law review articles referring to the Coase Theorem and 523 referring to the Prisoners' Dilemma.

10 The diffusion of the Coase Theorem has not been as broad as that of the Prisoners' Dilemma, though the recent award of the Nobel Prize to Coase may in time play a role in changing that. The Coase Theorem has been much discussed in law and, to some extent, in economics. *Cf.* Stewart Schwab, *Coase Defends Coase: Why Lawyers Listen and Economists Do Not,* 87 MICH. L. REV. 1171 (1989). However, the Theorem lacks the interdisciplinary sweep of the Dilemma, which has garnered attention from political scientists, psychologists, and biologists, as well as lawyers, economists, and philosophers. *See, e.g.,* AXELROD, *supra* note 5; ANATOL RAPOPORT & ALBERT M. CHAMMAH, PRISONER'S DILEMMA (1965) (experimental psychology); Robert Axelrod & William D. Hamilton, *The Evolution of Cooperation,* 211 SCI. 1390 (1981) (evolutionary biology); Robyn M. Dawes, *Social Dilemmas,* 31 ANN. REV. PSYCHOL. 169 (1980); Lumsden Malvern, *The Cyprus Conflict as a Prisoner's Dilemma Game,* 17 J. CONFLICT RESOL. 7 (1973); John Maynard Smith, *The Theory of Games and the Evolution of Animal Conflict,* 47 J. THEORETICAL BIOLOGY 209 (1974).

11 In a law review article, a certain pressure for normativity exists since the usual mode of legal writing involves normative argument, unlike the model prevailing in most other academic disciplines. The preference for explicit normativity in legal writing seems to me generally better than the preference for positive (or covertly normative) writing found in many other academic disciplines. *Cf.* Pierre Schlag, *Normative and Nowhere to Go,* 43 STAN. L. REV. 167 (1990); Symposium, *The Critique of Normativity,* 139 U. PA. L. REV. 801 (1991). Thus, because I think there is a legitimate "so what" issue about limiting this Article to a formal demonstration, I have made an effort to present the normative context in which the formal argument is situated.

Dilemma and Theorem. I will suggest that this choice between logistic stories such as the Dilemma and the Theorem and less logistic stories, though not one narrowly determined by left-right politics, is one that has political significance.[12]

Part I recounts the canonical Coase Theorem and Prisoners' Dilemma stories. Readers already familiar with the models may wish to skim this material or proceed directly to Part II, the central analytical section. There I show that the 2 × 2 matrix that formally represents the Coasean bargaining situation is a Prisoners' Dilemma matrix of a particular kind. In Part III I address the implications of the formal isomorphism between the Theorem and Dilemma for developing a critical perspective on the two models.

I. THE DILEMMA AND THE THEOREM, INFORMALLY

A.   The Prisoners' Dilemma[13]

You and another prisoner, your codefendant, are being held in separate cells, with no communication between you. The authorities have offered both of you a one-time-only plea bargain.[14] If you confess and incriminate your codefendant, you will receive a break on your sentence. How significant a break depends on whether your codefendant also confesses. If she does not, your testimony against her is valuable to the state in securing her conviction, and you will get only a one-year sentence. If she also confesses, however, your testimony is less valuable and you will both receive four years. If you do not confess and your codefendant does, you are in the worst position: you will receive a ten-year sentence. On the other hand, if you do not confess and your codefendant also does not confess, both of you are in relatively good shape: because the state's case without a confession is strong only for the lesser offenses you are charged with, you will both receive only two years.

Obviously, the best outcome for you and your codefendant is if neither of you confesses. However, a seemingly airtight argument for confessing exists. Your codefendant will either confess or not confess. If she con-

---

12   Others who agree with my formal point may find it conducive to moral and political purposes other than those I stress. I invite them to write their own articles. For examples of critical legal studies perspectives on law and economics consistent with the perspective of this article, see Thomas Heller, *The Importance of Normative Decision-Making: The Limitations of Legal Economics as a Basis for Liberal Jurisprudence—As Illustrated by the Regulation of Vacation Home Development*, 1976 Wis. L. Rev. 385; Mark Kelman, *Misunderstanding Social Life: A Critique of the Core Premises of "Law and Economics"*, 33 J. LEGAL EDUC. 274 (1983); Kelman, *supra* note 4; Duncan Kennedy, *Cost-Benefit Analysis of Entitlement Problems: A Critique*, 33 STAN. L. REV. 387 (1981); Duncan Kennedy & Frank Michelman, *Are Property and Contract Efficient?*, 8 HOFSTRA L. REV. 711 (1980); *see also* Baker, *supra* note 4; Herbert Hovenkamp, *Rationality in Law & Economics*, 60 GEO. WASH. L. REV. 293 (1992); Hovenkamp, *supra* note 4; Arthur Leff, *Economic Analysis of Law: Some Realism About Nominalism*, 60 VA. L. REV. 451 (1974).

13   "Prisoners' Dilemma" has been used here to express the shared nature of the problem facing the prisoners. The word "Prisoner's" is more widely used in the literature.

14   The Prisoners' Dilemma story assumes that the same official who is making the offer also has control over the sentence. An issue associated with the canonical Prisoners' Dilemma story (and not addressed in the Dilemma literature to my knowledge) is the morality of a prosecutor making the value of a confession contingent upon whether the other party confesses also. For a treatment of the morality of plea bargaining more generally, see Albert W. Alschuler, *The Changing Plea Bargaining Debate*, 69 CAL. L. REV. 652 (1981).

fesses, you are much better off if you also confess, because then you will receive four years instead of ten. If she does not confess, you are still better off if you confess, because then you will receive only one year instead of two. Thus, you should confess. Meanwhile, your codefendant, reasoning the same way, will also confess. As a consequence, you will both receive four years. But if the two of you avoided confessing, you would both be in a much better position, receiving two years instead of four. Thus the dilemma: you want the lighter sentence, but logic seems to condemn you and your codefendant to serving twice as long in prison.

B.   The Coase Theorem

A cattle rancher and a farmer are neighbors. The rancher's cattle occasionally stray onto the farmer's land and cause damage to the farmer's crops, for which the rancher is not legally liable. The annual damage caused to the farmer's crops by a cattle herd of the size that maximizes the rancher's income is fifty dollars.[15] This damage could be prevented if the rancher fenced her land, which she can do at an annual cost of fifteen dollars.[16]

Will the rancher fence her land? While fencing would maximize the total product of the farmer and the rancher, it might seem that the answer is no, since the rancher bears no liability for the damage her cattle cause to the farmer's crops. But that initial answer is incorrect. The rancher will fence her land, because the farmer will pay her to do it. The price is indeterminate, but both the farmer and the rancher have an incentive to make a deal. For any price above fifteen dollars, it is worthwhile to the rancher to put in the fence. For any price less than fifty dollars, it is worthwhile for the farmer to pay for the fence.

Now change the example by making the rancher legally liable for the damage his roving cattle cause. As in the earlier example, the rancher could prevent the fifty dollars of annual damage caused by the cattle if she installed a fence for fifteen dollars. But now assume the damage could also be prevented if the farmer, instead of planting wheat, planted sorghum. Because cattle are repulsed by sorghum, they will not approach it, much less trample it.[17] In this case, though, the farmer would suffer a loss of ten dollars in income compared to what he would make by planting wheat.

Will the farmer plant sorghum rather than wheat? As in the earlier example, the first impression is that he will not: while the farmer's and the rancher's total product would be greater if he did,[18] the fact that the rancher has to compensate the farmer for damaged wheat seemingly

---

15   In Coase's original story, the damage depends on the size of the herd: one steer causes one dollar of damage, two steers cause three dollars, three steers cause six dollars, and four steers cause $10 damage (the amount of damage rising more than proportionately to the number of steers could be justified because larger herds are more prone to stampede). Coase, *supra* note 2, at 3.

16   Assume no practical way exists (perhaps because of the topography) for the farmer to fence the property.

17   No claim is made that this part of my story bears any relationship to agricultural reality.

18   The total product would be greater by five dollars: if the farmer plants sorghum, which costs him $10, the rancher will not have to fence, which costs her $15.

means that wheat will be planted. But for exactly the same reason stated earlier, the first impression is wrong. Sorghum will be planted because the rancher will pay the farmer to do so. The rancher has an incentive to pay up to fifteen dollars, and the farmer has an incentive to accept anything over ten dollars.

In both examples, then, one notes that the optimal result is achieved, regardless of the liability rule in effect. The casual impression that social efficiency[19] depends on the imposition of proper regulations or liability rules turns out to be incorrect. Through their own bargaining, the parties are entirely capable of reaching an efficient outcome for themselves.

II.   THE DILEMMA AND THE THEOREM, FORMALLY

A.   The Prisoners' Dilemma

The Dilemma was early on stated as, and has since been viewed as, a formal problem[20] in game theory, expressed in the form of a 2 × 2 matrix.[21] One standard formulation of the Dilemma as a matrix is given below:[22]

THE PRISONERS' DILEMMA

| | | Column Player | |
		Strategy 1	Strategy 2
	Strategy 1	(1, 1)	(5, 0)
Row Player	Strategy 2	(0, 5)	(3, 3)

Note: The payoffs for Row Player are listed first in each box and those for Column Player are listed second. Assume that higher payoffs are better than lower ones.

---

19   I define social efficiency as the sum of the production of the two parties. Coase's own exposition of the rancher-farmer story is considerably longer, employs more economic terminology (although with the clarity that is a signal virtue in Coase's writing), and is more detailed than my version. *See* Coase, *supra* note 2, at 2-8. Nevertheless, the Coase Theorem story presented here is thoroughly consistent with that presented in *The Problem of Social Cost, id.*

20   Two RAND researchers, Merrill Flood and Melvin Dresher, are credited with the discovery of the Dilemma in 1950. Albert Tucker provided the story. HOFSTADTER, *supra* note 5, at 715. In 1944, John Von Neumann and Oskar Morgenstern set out a formal, mathematical approach to game theory that was the intellectual background to the formalized statement of the Prisoners' Dilemma as a 2 × 2 matrix. JOHN VON NEUMANN & OSKAR MORGENSTERN, THEORY OF GAMES AND ECONOMIC BEHAVIOR (1953).

21   While this section follows the convention of expressing the Dilemma as a 2 × 2 matrix, it is also perfectly possible to represent the Dilemma in an extensive or game tree format. *See* DAVID M. KREPS, GAME THEORY AND ECONOMIC MODELING (1990) (underscoring the logical identity between matrix and extensive forms of representation); Ian Ayres, *Playing Games with the Law*, 42 STAN. L. REV. 1291 (1990) (reviewing ERIC RASMUSEN, GAMES AND INFORMATION (1990)). Such a format can have advantages in representing a game in which the parties move sequentially rather than simultaneously. Nonetheless, since the Dilemma involves simultaneous movement, the matrix format serves its present purpose.

22   This formulation is similar to that in AXELROD, *supra* note 5, although Axelrod introduces descriptive labels to the strategies and the payoffs. Axelrod calls strategy 1 "cooperation" and strategy 2 "defection"; the 1 payoff is the "punishment for defection," the 3 payoff is the "reward for cooperation," the 5 payoff is the "temptation," and the 0 payoff is the "sucker's payoff." *Id.* at 8. I have eschewed these evocative labels to avoid characterizing the matrix in a way that limits it to one story or kind of story.

In the Prisoners' Dilemma matrix, both players have an incentive to agree to play strategy 2, so that they can each get 3. On the other hand, both individuals also have an incentive to play strategy 1, reasoning that no matter what the other player does, one is better off with strategy 1 (getting 5 instead of 3 if the other plays strategy 2, and 1 instead of 0 if the other plays strategy 1).[23]

Of course, the Prisoners' Dilemma matrix does not need to use the particular numbers in the matrix shown. All that is necessary to produce a Dilemma is a certain relationship between the numbers. For both the Row Player and the Column Player, the payoff A,[24] for playing strategy 1 when the other player plays strategy 2, has to be greater than B, the payoff for playing strategy 2 when the other player plays strategy 2. B in turn has to be greater than C, the payoff for playing strategy 1 when the other player plays strategy 1, and C has to be greater than D, the payoff for playing strategy 2 when the other player plays strategy 1.[25] Or, $A > B > C > D$.[26]

---

23   In the language of game theory, strategy 1 strongly dominates strategy 2. In other words, the payoffs for choosing strategy 1 are greater than those for choosing strategy 2, regardless of what state of the world prevails or what the other player chooses to do. A strategy weakly dominates another when the payoffs for it are greater than or equal to those for the other strategy regardless of what the other player does. See Robert Nozick, *Newcomb's Problem and Two Principles of Choice, in* ESSAYS IN HONOR OF CARL G. HEMPEL (Nicholas Rescher ed., 1970) for a reflection on the dominance principle and its limitations (although Nozick's conclusion that the dominance principle should dominate in deciding what one does in Newcomb's Problem seems dubious).

24   The usage here follows AXELROD, *supra* note 5, at 8.

25   While the typical formulation of the Dilemma has symmetrical payoffs for both players, that is not necessary as long as the relationships between each player's A, B, C, and D payoffs are as described.

26   Another condition occasionally seen as necessary to define a Prisoner's Dilemma matrix is that the "temptation payoff" plus the "sucker's payoff" is less than twice the "reward for cooperation." *See supra* note 22 (explaining the descriptive labels); AXELROD, *supra* note 5, at 206. The point of this condition is that it ensures that both players' playing strategy 2 results in the highest total payoff, higher than the total obtained by playing strategy 1 while the other plays strategy 2— that is, "mutual cooperation" works better than "alternating exploitation." For present purposes, I do not believe this condition adds anything necessary to the formal statement of the Dilemma, and it is not included here as a condition. For further explanation of this point, see *infra* note 33.

One can think of various ways in which the force of the argument for playing strategy 1 in the Dilemma matrix is weakened. If one is going to play the Dilemma again with the same player, or for that matter with someone else who will be aware of your conduct in this game, the case for choosing strategy 1 weakens since playing strategy 2 may induce the other player to do so in future games. On this point, however, the caveat should be noted that if both players know the number of trials of the game in advance, there is an argument, accepted by most but not all commentators on the Dilemma, that it is to one's advantage to choose only strategy 1: on the last round, you and the other individual will play strategy 1—so, knowing that, you and the other will also play strategy 1 on the next to last round, and the next to next to last, and so on all the way up to and including the first round. *See* LUCE & RAIFFA, *supra* note 5, at 98-99; J. Howard Sobel, *The Need for Coercion, in* COERCION: NOMOS XIV (J.R. Pennock & J.W. Chapman eds., 1972). *But see* RUSSELL HARDIN, COLLECTIVE ACTION 146 (1982).

In stating the Prisoners' Dilemma, it is sometimes the practice to qualify the matrix with provisos, such as "no communication between players," "egoistic motives only," "one trial only," or "no enforceable agreements." These provisos are designed to highlight the logical case for choosing strategy 1 and to negate the case for choosing strategy 2. For purposes of this paper, a Dilemma exists on the basis of the formal relationship of the payoff figures in the matrix; that is because the provisos, while undoubtedly relevant to the player's decision-making process, do not define the Dilemma matrix itself.

B.   The Coase Theorem

To formalize the logic of Coasean bargaining and to compare it with that of the Dilemma, it is first necessary to put Coasean bargaining in the game-theoretic language of a payoff matrix. Fortunately, although Coase did not describe what he was doing as game theory, the Coasean bargaining that is the Theorem's central feature readily lends itself to formulation as a game-theoretic proposition.[27]

The initial or naïve understanding of the situation faced by the rancher and the farmer (or any two parties, one of whose actions affects the other) can be put in matrix form as follows:

THE STARTING POINT OF THE COASEAN STORY

Column Player (Farmer)

Row Player	Strategy 1	(2, 1)
(Rancher)	Strategy 2	(0, 5)

Note: Row Player's payoffs are shown before Column Player's payoffs in both boxes. Assume both players seek to maximize their payoffs.

In this simple 2 × 1 matrix, only Row Player has a choice. She will choose strategy 1 (for instance, by not fencing her land, or, in the second example, by growing wheat rather than sorghum). However, the basic Coasean point, translated into game theoretic terms, is that this 2 × 1 ma-

---

27   For a more comprehensive discussion of Coase's logic as a proposition in game theory, see Wayne Eastman, *Everything's Up for Grabs: The Coasean Story in Game-Theoretic Terms*, 31 NEW ENG. L. REV. (forthcoming 1996). Coasean logic can be formalized in a non-game-theoretic fashion. *See* STIGLER, *supra* note 3; Hovenkamp, *supra* note 4; Hovenkamp, *supra* note 12; Donald H. Regan, *The Problem of Social Cost Revisited*, 15 J. L. & ECON. 427 (1972). A comparison of the formalization carried out in the text with alternative formalizations is a project beyond my immediate scope, though I would suggest that a game-theoretic formalization preserves the flavor of Coase's stories better than the alternative formalizations.

In addition, there is a vast amount of literature that inquires into various aspects of the Theorem with which I am not concerned here, such as the distinction between the Coasean claim that an efficient outcome will be reached regardless of the liability rule (the efficiency claim) and the Coasean claim that the same efficient result will be reached regardless of the rule (the invariance claim). *See, e.g.,* Calabresi, *supra* note 4 (rejecting the argument that long-run wealth effects undermine the Theorem); Robert Cooter, *The Cost of Coase*, 11 J. LEGAL. STUD. 1 (1987) (accepting the efficiency claim, and rejecting the invariance claim); Elizabeth Hoffman & Matthew L. Spitzer, *The Coase Theorem: Some Experimental Tests*, 25 J. L. & ECON. 73 (1982); G. Warren Nutter, *The Coase Theorem on Social Cost: A Footnote*, 11 J. L. & ECON. 503 (1968) (reaching a conclusion similar to Calabresi's); Regan, *supra* (stating and criticizing the efficiency and invariance claims).

The focus in this Article on interpreting the Coase Theorem in game theoretic terms is not a new one. However, to my knowledge, this Article's thesis regarding the formal equivalence of Coasean bargaining and the Dilemma has not been articulated. For instance, in his article Regan asserts that the Coase Theorem is a proposition in the theory of games. *Id.* at 428. Similarly, in his articles on the Coase Theorem, Hovenkamp deals with strategic bargaining issues that arise when the parties are negotiating about how to divide a surplus, and discusses the Prisoners' Dilemma. Hovenkamp, *supra* note 4, at 787-91; *see also* Hovenkamp, *supra* note 12, at 311-12. But neither author makes the argument in this Article about the formal connection between the Dilemma and Coasean bargaining.

trix does not accurately reflect the actual decision situation faced by the two parties. In fact, Column Player and Row Player have a choice. Column Player can choose to pay Row Player a certain amount pursuant to an agreement that Row Player choose strategy 2 instead of strategy 1. Row Player, as in the earlier example, has a choice whether to play strategy 1 or strategy 2.[28] More generally, the Coasean point is that a game theory matrix showing fixed payoffs for the players is always subject to alteration, based on the players' agreement to redistribute the payoffs.[29] Thus, the situation facing the rancher and the farmer is properly represented as a 2 × 2 rather than a 2 × 1 matrix:

CoASEAN BARGAINING

Column Player (Farmer)

		Strategy 1	Strategy 2
Row Player	Strategy 1	(2, 1)	(2+x, 1−x)
(Rancher)	Strategy 2	(0, 5)	(x, 5−x)

Note: The payoff amounts for the boxes in the right hand column derive from Column Player agreeing to pay x to Row Player in return for Row's agreement to play strategy 2. The amount x is indeterminate; in the example given, Row and Column might agree on any amount between 2 and 4.

In this 2 × 2 matrix, both players will be better off if they make and then comply with an agreement to play strategy 2 (if the rancher fences and the farmer pays the promised amount for the fence) than if they fail to reach an agreement or to comply with it. But both players have an incentive to threaten not to reach an agreement unless it is one favorable to them, and to provide arguments designed to make their threats credible. For example, the farmer might say that for the rancher to receive anything more than the actual cost of 2 for fencing would be to reward the antisocial marauding of his cattle in a way that the farmer cannot countenance. After all, in the absence of the rancher she would receive 5, while in her absence the rancher would receive only 2. On the other side, the rancher might say that for him to receive anything less than 3 from the farmer would be unfair, since that would mean that the original differential in favor of the rancher in the situation in which there is no fencing would be reduced in favor of the farmer.[30]

---

28   Row Player has a choice whether to play strategy 2 according to the agreement or to renege by playing strategy 1.
29   In Eastman, *supra* note 27, I argue that this ability of the parties to change their payoffs through agreement is the key Coasean twist on standard game theory. To relate that argument to the one made in this Article: the logic of the Coase Theorem is the logic of payoff mutability, while the logic of Coasean bargaining—the subject here—is the logic of the Prisoners' Dilemma.
30   For articles discussing incentives for strategic bargaining that may lead to an agreement not being reached, see Cooter, *supra* note 27; Robert Cooter, Stephen Marks & Robert Mnookin, *Bargaining in the Shadow of the Law: A Testable Model of Strategic Behavior*, 11 J. LEGAL STUD. 225 (1982); Eastman, *supra* note 27.

Assuming they do reach an agreement, both players have an incentive to break it by playing strategy 1. The farmer can reason as follows: "Either the rancher will keep his promise to fence or he will not. If he does not keep his promise, I am better off breaking my promise; and if keeps his promise, I am also better off breaking my promise."[31] If circumstances can be arranged so that the exchange can be made at one time, so that fencing takes place simultaneously with the farmer's payment, or if there is an effective contract law system that allows both parties to sue for breach of promise, the incentives both parties have to break any commitment they make can be countered. But both sides have an incentive to violate any commitment they make.

C. Coasean Bargaining Involves a Dilemma

As just described, the situation faced by Coase's farmer and rancher sounds like a Prisoners' Dilemma. That is in fact the case. The Coase Theorem story of bargaining parties translates into a Prisoners' Dilemma no matter what numbers one chooses for the initial 2×1 matrix (as long as those numbers are such that Coasean bargaining is warranted).[32] The Coasean bargaining version of the Prisoners' Dilemma matrix is defined by the conditions that for both players the payoff A for reneging on a promise while the other player complies is greater than payoff B for complying when the other player complies, which in turn is greater than payoff C for reneging when the other player reneges, which in turn is greater than payoff D for complying when the other player reneges. Or, as in the general Prisoners' Dilemma, $A > B > C > D$.[33] The conclusion is that the matrix representing Coasean bargaining is a Prisoners' Dilemma matrix.[34]

---

31 The reasoning for the rancher is parallel: "Either the farmer will keep his promise or he will not. If he does not, I am better off not keeping mine; and if he does, I am also better off not keeping mine."

32 Coasean bargaining is warranted whenever the loss incurred by the affected party (Column Player) is greater than the gain received by the affecting party (Row Player). That is, if we label the payoffs in the 2×1 matrix as (C1, C2) for the upper box and (D1, A2) for the lower box, $A2 - C2 > C1 - D1$.

33 While the point is not a crucial one to this Article's argument, the Coase Theorem matrix is properly described as a subset of the Prisoners' Dilemma matrix rather than vice versa for two reasons. First, the values the payoffs can take in the Coase Theorem version of the Dilemma, although indeterminate, are more restricted than those in the general version of the Dilemma. Specifically, the right column boxes in the 2×2 Coase Theorem matrix are most plausibly represented as equal to the sum of the payoffs in the corresponding left column boxes—that is, if the upper left box payoffs are (1, 1), the zero-sum transfer involved in Column's paying Row a certain amount x is plausibly represented by the payoffs in the upper right box being $(1 + x, 1 - x)$. No such restrictive condition applies in the general form of the Dilemma matrix. Second, while in the general version of the Dilemma it is not necessary that the players' payoffs be regarded as commensurable, that assumption is necessary to allow Coasean bargaining to work. In other words, the Coase Theorem assumes a producer rather than a consumer orientation. *See* Baker, *supra* note 4; Kelman, *supra* note 4. The Dilemma, on the other hand, is compatible with either a producer or a consumer orientation.

34 Because in the Coase Theorem case, playing strategy 2 for both players yields a total payoff that is equal to, rather than greater than the payoff for Column Player playing strategy 1 (breaking the promise by failing to pay) while Row Player plays strategy 2 (keeping the promise), the sometimes stated condition for a Dilemma, *see supra* note 26, that the total of the strategy 2 payoffs exceed the total of the payoffs for any other box is not met in the Coase Theorem case. That does not mean, however, that the Coasean bargaining should not be considered a Prisoners' Dilemma. The choice situation in Coasean bargaining has all the essential features of the Di-

It is perfectly correct to point out that the Coase Theorem and the Prisoners' Dilemma, as often presented and understood, involve conflicting assumptions. One might counter my claim of formal identity between Coasean bargaining and the Dilemma by noting the contradiction between the assumption of no transactions costs, made by Coase and most subsequent exponents of the Theorem,[35] and the assumption that no binding agreements are possible, made by some exponents of the Dilemma.[36] These two assumptions are indeed contradictory. That does not, however, negate the point that there is a formal connection between the Theorem and the Dilemma. The matrix representing Coasean bargaining is still a Dilemma matrix.

The difference between Coasean bargaining and the Dilemma lies in assumptions, intuitions, or judgments about the world and about human behavior that are expressed in provisos to the matrices, not in the formal structures of the matrices, which are identical. There may be a plausible practical basis in some cases for distinguishing between a "Coase Theorem" situation, in which transaction costs are low and the parties will bargain to an efficient resolution, and a "Prisoners' Dilemma" situation, in which parties who fail to cooperate will not bargain to an efficient resolution. But the difference between the "Coasean bargaining" and "Prisoners' Dilemma" situations is an empirical, contextual one, not one inhering in their formal structures. Logic does not carry the day, for the logics of the two situations are the same.

III. WHAT IT MEANS

The formal identity between Coasean bargaining and the Prisoners' Dilemma does not have determinate political or moral implications. Nonetheless, recognition of the formal identity can be conducive to developing a critical perspective on the way the Theorem and the Dilemma are typically presented. In this section I will briefly discuss how recognizing the formal identity of the two models can be helpful in that regard.

The Coase Theorem and the Prisoners' Dilemma are stories about the world, just as *Madame Bovary* and the account of Moses in Sinai are stories. The manner in which the Coase Theorem and Prisoners' Dilemma stories

---

lemma. In particular, it is not in the interests of the players to bargain for a "mutual exploitation" situation in which Column and Row alternate in playing strategy 1 and strategy 2. Here, the condition that the total payoffs from both playing strategy 2 exceed the total from any other box is not necessary to define a Dilemma. Further, in an iterated Coase Theorem, one could lower the payoffs for Row Player and Column Player in the lower left-hand box to reflect the future negative effects of promise-breaking; then the (minor) problem of the lower right box payoff total not exceeding the lower left payoff total no longer exists.

The reader might wonder whether most 2 × 2 matrices fit the conditions for a Dilemma. The answer is no. If payoff numbers for a 2 × 2 matrix are chosen at random, the result is unlikely to be a Dilemma. If one assumes symmetrical payoffs for Row and Column, the chance that randomly chosen numbers will define a Dilemma is 1 in 4 or less, depending upon the assumptions made about the range from which one chooses and the possibility of duplication. If one assumes no relationship between Row's and Column's payoffs, the chance that a Dilemma will occur at random is 1 in 32 or less.

35 *See* Coase, *supra* note 2; STIGLER, *supra* note 3.

36 *See, e.g.*, Richmond Campbell, *Background for the Uninitiated, in* PARADOXES OF RATIONALITY AND COOPERATION, *supra* note 5, at 3, 8.

are told, however, is rather distinctive. The characters in the Dilemma and the Theorem are abstract, generic beings, and their fate is presented not as the outcome of their particular qualities of character and temperament but as a working out of logic, which works through them as God works through Moses.

In their fusion of formal logic with particular morally and politically significant stories,[37] the Coase Theorem and the Prisoners' Dilemma are problematic. The problem is not that it is inadmissible to tell a story in a way that enlists logic; to lay down such a canon would be an unwarranted restriction of the storyteller's art. The problem is rather one of authority. In lay understanding, logic connotes certainty.[38] To tell a story that fuses powerful political and moral messages[39] with a formal logical structure, as both the Coase Theorem and the Prisoners' Dilemma do, is to imply that the particular political and moral content of the story one tells has the special reliability of logic.[40]

---

37   The aim here is not to imply a rigid dichotomy between "logic" and "story." The 2 × 2 matrix I refer to in this Article, which embodies the formal logic of the Dilemma and the Theorem, is itself a particular device for telling stories, and one might reasonably say that logical truisms such as the dominance principle are also such devices. The question about the conceptual distinctions between logic and story, and logic and moral/political message, concerns whether these distinctions are valuable for certain purposes. This Article is not concerned with whether these distinctions are true in the sense that they correspond to an underlying reality. *See generally* WILLIAM JAMES, ESSAYS IN PRAGMATISM (1948); FRIEDRICH NIETZSCHE, BEYOND GOOD AND EVIL (1886), *reprinted in* BASIC WRITINGS OF NIETZSCHE (Walter Kaufmann, ed. & trans., 1968).

38   For a non-lay perspective on logic, which considers the possibility of the abandonment of logical propositions, see WILLARD VAN ORMAND QUINE, FROM A LOGICAL POINT OF VIEW (2d ed. 1980).

39   While the Theorem and the Dilemma stories both have standard messages, the subtlety and internal contradiction within these messages, as well as the possibility for the transmutation of the standard messages into different forms should be appreciated. *See, e.g.,* Hovenkamp, *supra* note 4; Schlag, *supra* note 4. The standard Coasean message that people can bargain efficiently through side payments carries within it the standard counter, or shadow message, that intervention makes sense because of the pervasiveness and significance of transaction costs. *See, e.g.,* Calabresi, *supra* note 4. The standard Prisoners' Dilemma message that self-interest fails to achieve optimum results carries within itself the standard counter-message that, with communication and a contract law system, self-interest will achieve optimum results. This politically-charged indeterminacy is one reason why the Theorem and the Dilemma are indeed good stories—and the critical point here should not be taken to deny this valuable aspect of the Theorem and the Dilemma. *See* Eastman, *supra* note 1. At the same time, one should also note that the ability of the models to inculcate mainstream opinion is enhanced by the fashion in which they define opposing positions that can both be comfortably embraced within the political mainstream. The opposition in the Dilemma between a hard-edged pursuit of self-interest and a softer one is naturally, though not ineluctably, identified with the familiar opposition between mainstream free-market conservativism and a milder mainstream liberalism. The opposition in the Coase Theorem between high and low transaction cost situations is naturally identified with the familiar mainstream debate over the areas of competency of government and markets.

40   In practice, the use of form to debunk formalism in economic models may well have political consequences, though such consequences are not assured, especially given that one of the models discussed here, the Dilemma, is a model heavily relied upon by political liberals. The claim that those inclined toward open-ended legal standards have an affinity for altruism, and those inclined toward formally realizable rules have an affinity for individualism, has played a prominent role in critical legal studies. *See generally* MARK KELMAN, A GUIDE TO CRITICAL LEGAL STUDIES (1987); Duncan Kennedy, *Form and Substance in Private Law Adjudication*, 89 HARV. L. REV. 1685 (1976). Similarly, it is plausible—though here I will only suggest the point rather than develop a case for it—that those especially drawn to repose faith in economic models tend to be drawn to conservative, individualistic politics, while those who believe little in such models tend to be drawn to liberal politics. If that is the case, the debunking project of this Article is not a

The analysis of form I have undertaken is a way to unmask formalism. Formally, Coasean bargaining is represented by a Prisoners' Dilemma. There is a strong difference between the standard anti-interventionist Coasean message and the standard interventionist Prisoners' Dilemma message[41] which is grounded not in formal logic, but in assumptions, judgments, intuitions, and evaluations about the world and about human nature.

If Coasean bargaining and the Dilemma are presented and considered separately as entirely distinct models, as they generally are,[42] both models gain a certain power and authority from the implication that their different morally and politically-charged stories are grounded in logic. When the form and substance of the models are considered together, however, this distinctive formalist power and authority is no longer plausible. If the formal structure of Coasean bargaining and the Dilemma were different, one might argue that the morally and politically-charged message of Coasean bargaining is grounded in logical point A and the different message of the Dilemma is grounded in logical point B.[43] If the formal representations of the two situations are identical, as they in fact are, one cannot argue that the different messages are thus grounded in logic. One is left to see the stories of Coasean bargaining and the Prisoners' Dilemma as stories like other stories, rather than as stories whose messages have a special logical imprimatur.[44]

---

politically neutral one, but one that may be of more value to those drawn to liberal politics than those drawn to conservative individualism.

41 This point about the difference between the Coasean and Prisoners' Dilemma messages seems clearly true even allowing, as stressed before, the indeterminacy and multi-facetedness of both models' messages. *See supra* note 39.

42 See KELMAN, *supra* note 40, at 114-85 for an argument that mainstream legal (and law and economics) discourse represses awareness of contradiction. That claim is correct, I believe, in regard to the simultaneous ascendancy of contrasting Coasean bargaining and Prisoners' Dilemma perspectives on cooperation by self-interested parties—an ascendancy that relies on splitting the two models into separate mental boxes rather than considering them together.

43 Such an argument would still be wrong, I believe, but that issue is beyond the scope of this Article.

44 One might wonder why, from a critical point of view, one should bother with the specific analysis of form undertaken here. Instead, why not make or refer to a general argument against formalism? For example, instead of examining the Theorem and the Dilemma, why not argue that logical concepts do not resolve substantive moral and political questions? A major problem is that such an argument is at too high a level of abstraction to have much practical effect. In the modern legal academy, scholars from the right, left, and center proclaim themselves anti-formalists. *See, e.g.,* RICHARD A. POSNER, THE PROBLEMS OF JURISPRUDENCE (1990); Richard A. Posner, *The Jurisprudence of Skepticism,* 86 MICH. L. REV. 827 (1988). Thus, it might seem that formalism is a merely superficial or vestigial phenomenon that can be treated as trivial and dismissed with a cursory general argument.

I believe that assertion is entirely wrong. Formalism is a deep-rooted, if often disavowed, way of thinking and feeling that reappears constantly in spite of ourselves in the way we understand and create the world legally. If that is so, an effort to avoid formalism by sprinkling about the holy water of anti-formalism will not likely achieve its intended result. One can do better by combining the anti-formalist message (a message that is better conveyed in moral and political terms than in epistemological ones) with a careful critical analysis of particular forms. Specifically, the formalism associated with the Coase Theorem and the Prisoners' Dilemma can be understood and criticized effectively through analyzing their formal structures. A more general anti-formalist argument is also available to criticize the political use of the Dilemma and the Theorem. My claim here, though, is that such a broad anti-formalist argument is likely to be less plausible and to work less well than a critical analysis of the specific forms of the models.

One is remitted, as one should be, to the messier domains of empiricism, morality, and politics.[45]

---

45 Without in any way revoking the critical point made in the text, I note that I do not intend to imply that the Dilemma and the Theorem, or mainstream economic models more generally, are without value. Critical law and economics work that focuses on formalism in economic models needs model-building law and economics, not only in the straightforward way in which Washington lawyers who oppose government regulations need regulations, but also in the trickier ways in which opposition needs a mainstream position against which to define itself and from which to draw its own affirmative agenda. Similarly, mainstream law and economics needs critical law and economics. Critics of formalism in law and economics are doing work that mainstream scholars, most of whom are at least nominally committed to anti-formalism, should want to have done, and done not only for reasons of intellectual integrity. Although criticism may rankle, law and economics is more viable as an intellectual project to the extent that it includes nay-saying as well as aye-saying advocacy. Of course, the political and moral divisions that may separate critical and mainstream law and economics scholars should not be denied. Neither should it be denied that they are, perhaps in spite of themselves, engaged in a symbiotic enterprise. For an argument that critics of prevailing stories in law and economics should participate in the discourse of the discipline by creating alternative stories, see Eastman, *supra* note 1.