

LEGAL ANALYSIS OF ECONOMICS: SOLVING THE PROBLEM OF RATIONAL COMMITMENT

BRUCE CHAPMAN*

I. LAW AND ECONOMICS AS A MORE EQUAL PARTNERSHIP

Lawyers and economists meet under either of two banners that signal their joint movement. One of these is “law and economics”; the other is “economic analysis of law.”¹ While the two labels are used more or less interchangeably, there is an important difference in emphasis between them. “Law and economics” suggests the possibility of a more equal partnership; we can imagine that either term in the pair might have gone first, and we are none the wiser, on observing a particular order in the terms, as to what role either law or economics plays in the partnership. But “economic analysis of law” is a good deal less ambiguous. It suggests something that has largely been true about the interchange between law and economics, namely, that economics provides the method of analysis and law the subject matter to which this analysis is applied.²

Of course, in one sense that this is the truth of the matter about law and economics should not be surprising. Over the last four decades we have become used to economics being characterized less by a distinctive subject matter and more by its particular analytical approach to social behavior. The time has long past when economics

* Faculty of Law, University of Toronto. I am grateful to Shachar Lifshitz, Joe Mintoff, Oren Perez, and, particularly, Wlodek Rabinowicz and Horacio Spector for very helpful comments on an earlier version of the arguments that appear in this paper. The comments of the various participants in the Special Workshop on Law and Economics and Legal Scholarship, 21st IVR World Congress, Lund, Sweden (August 12–18, 2003), as well as those in the Annual Meeting of the Canadian Law and Economics Association in Toronto (September 2003) are also gratefully acknowledged. Funding for the research on this paper was provided by the Social Science and Humanities Research Council of Canada.

1. The names of the different introductory texts in the area illustrate this point. Compare RICHARD A. POSNER, *ECONOMIC ANALYSIS OF LAW* (6th ed. 2003), with A. MITCHELL POLINSKY, *AN INTRODUCTION TO LAW AND ECONOMICS* (2d ed. 1989), and ROBERT COOTER & THOMAS ULEN, *LAW AND ECONOMICS* (3d ed. 2000).

2. See Richard A. Posner, *Law and Economics in Common Law and Civil Law Nations*, 7 *ASS'NS: J. LEGAL & SOC. THEORY* 77, 77 (2003) (characterizing the banner “law and economics” as misleading).

confined its gaze to markets or even institutions of market regulation or taxation. There are now, in addition to the economic analysis of law, economic analyses of political processes, voting behavior, bureaucracies, family life, social norms (including socially abnormal behavior such as suicide), religious practices, and biological systems.³ In the face of this general theoretical onslaught upon such a broad range of subject matters, it would be more of a surprise and, perhaps, even something of a slight, if law had not drawn the attention of economics.

Nevertheless, in another respect there is something odd about the strong asymmetry of the partnership. What is definitive of the economic approach to all these different subject matters is the use of the rational actor model. The constant assumption across all the different institutions that are analyzed is that these institutions are populated by *homo economicus*, the sort of being who can and does (among other things⁴) rationally order any set of alternatives from best to worst (according to whatever criteria this being happens to value) and consistently chooses that alternative which is best according to this overall preference ordering. Yet law and legal theory can *also* claim to offer an account of rational decision-making. Indeed, the exploration and articulation of the nature of legal reasoning has long been one of the central projects of legal theorists.⁵ And so it is a little surprising, to say the least, that the economist can so successfully lay siege to the edifice of law without having to engage the legal theorist, at least occasionally, on the rationality front.⁶

3. For a good sense of the range of topics that are discussed, see GARY S. BECKER, *THE ECONOMIC APPROACH TO HUMAN BEHAVIOR* (1976).

4. Where the individual faces uncertainty, a slightly more elaborate set of axioms is required to rationalize preference or utility maximizing choice. For a classic, and clear, articulation of the required axioms, see R. DUNCAN LUCE & HOWARD RAIFFA, *GAMES AND DECISIONS: INTRODUCTION AND CRITICAL SURVEY* 19–31 (1957).

5. See, e.g., STEVEN J. BURTON, *AN INTRODUCTION TO LAW AND LEGAL REASONING* (1985); JAAP C. HAGE, *REASONING WITH RULES: AN ESSAY ON LEGAL REASONING AND ITS UNDERLYING LOGIC* 9 (1997); NEIL MACCORMICK, *LEGAL REASONING AND LEGAL THEORY* (1978); CASS R. SUNSTEIN, *LEGAL REASONING AND POLITICAL CONFLICT* (1996).

6. It is not as if economics is problem-free on this front either. Recent work by behavioral psychologists (and experimental economists) has raised serious questions about whether it is appropriate to assume that agents are economically rational as a matter of fact. Experiments appear to show that we are neither as “good” (e.g., as competent with probability theory or as consistent in our choices over time) nor as “bad” (e.g., as self-interested or uninterested in fairness) as the economic model of the rational actor predicts. Christine Jolls, Cass R. Sunstein, & Richard H. Thaler, *A Behavioral Approach to Law and Economics*, in *BEHAVIORAL LAW AND ECONOMICS* 21–26 (Cass R. Sunstein ed., 2000); Russell Korobkin, *Behavioral Economics, Contract Formation, and Contract Law*, in *BEHAVIORAL LAW AND ECONOMICS*, *supra*, at 116–

This Article begins such an engagement. In the interest of developing a more equal partnership between the two traditions of decision theory that have separately grown up within economics and the law, we might think of the Article as offering a countervailing “legal analysis of economics” to the prevailing “economic analysis of law.” To execute this plan, I need to deliver on three separate fronts. First, I need to identify a peculiarly economic problem, or subject matter, to which the theoretical legal analysis will be applied. Not surprisingly, given that economics is now defined more as an analytical method than as a subject matter, the economic problem up for legal analysis will itself be a theoretical one, although the economist will claim, surely, that it has its real world applications. I have chosen to look into the problem of rational commitment and set this problem out in some detail in Part II. Second, in the interest of perfect symmetry, or the development of a mirror image legal analysis of economics, I need to provide an alternative rational actor model as the focal point of the analysis I will apply to the economic problem. Using some recent work by the philosopher-economist John Broome, I will develop this alternative account of rational behavior in Part III.⁷ Finally, in Part IV, I argue that this alternative rational actor model is legal in that it is manifested in common law adjudication and, more particularly, in the special relationship that exists between cases and defeasible legal rules. As we shall see, under the more complex rationality that is exemplified by legal reasoning, legal rules both determine the results of particular cases and in turn are determined by them. This, I argue, is exactly the sort of rationality that the economist needs to solve the problem of rational commitment. Lastly, I provide some brief concluding remarks in Part V.

42. This volume, *BEHAVIORAL LAW AND ECONOMICS*, *supra*, contains an excellent collection of papers in this area.

7. See John Broome, *Normative Requirements*, 12 *RATIO* 398 (1999) [hereinafter Broome 1]; John Broome, *Are Intentions Reasons? And How Should We Cope with Incommensurable Values?*, in *PRACTICAL RATIONALITY AND PREFERENCE: ESSAYS FOR DAVID GAUTHIER* 98, 100 (Christopher W. Morris & Arthur Ripstein eds., 2001) [hereinafter Broome 2]; John Broome & Christian Piller, *Normative Practical Reasoning*, 75 (Supp.) *PROC. ARISTOTELIAN SOC'Y FOR SYSTEMATIC STUDY PHIL.* 175–76 (2001) [hereinafter Broome 3]; John Broome, *Practical Reasoning*, in *REASON AND NATURE: ESSAYS IN THE THEORY OF RATIONALITY* 85, 86 (José Luis Bermúdez & Alan Millar eds., 2002) [hereinafter Broome 4]; and John Broome, *Reasons*, in *REASON AND VALUE: ESSAYS ON THE MORAL PHILOSOPHY OF JOSEPH RAZ* (R. Jay Wallace et al. eds., forthcoming 2004) [hereinafter Broome 5]. All subsequent page references to the forthcoming article Broome 5 is to a type scripted version of the article that is on file with the author and is available at <http://users.ox.ac.uk/~sfop0060/framepublication-ions.shtml>.

II. THE PROBLEM OF RATIONAL COMMITMENT

It is odd that the economic theory of rational choice has so little to say about the role of reasons in rational decision-making.⁸ One would have thought that the distinguishing feature of rational choice would be that it involves reasons. However, the economist might plausibly reply that there is nothing in the formal structure of rational choice theory that *precludes* any particular reason from influencing some choice between two alternatives x and y . These different reasons (which, it should be emphasized, may be self-interested or other-regarding, consequentialist or deontological, or objective or subjective) are simply taken as given. Moreover, the economist might well argue that the minimal rationality conditions that are required of an agent, for example, that her preferences be transitive, are only required so that she can satisfy her various reasons in the choices she makes. For if an agent, for whatever reason, preferred x to y , y to z , and z to x , in violation of transitivity, then it would not be possible for her to choose any one of these three alternatives without choosing contrary to some preference or the requirements of some reason. Thus, in requiring certain formal conditions of rationality, such as transitivity of preferences, the economist is really only trying to meet the same fundamental concern identified by many theorists as essential to rationality, namely, that in every possible choice a rational agent must always act “only and always for undefeated reasons.”⁹ Reasons give rise to a preference for doing x rather than y , and rationality consists in following that preference. It is irrational, in other words, to act contrary to a preference, or contrary to the reason that lies behind it.

All this seems quite compelling and so it is somewhat surprising to learn how quickly it can generate a problem. The difficulty is that an agent can have a reason, or a preference, to choose to do something that this same agent has no reason actually to do, or which will

8. For observations that this is a fair characterization of rational choice theory, see Paul Anand, *The Nature of Rational Choice and The Foundations of Statistics*, 43 OXFORD ECON. PAPERS 199, 210 (1991); and Amartya Sen, *Rationality and Uncertainty*, 18 THEORY & DECISION 109, 110–11 (1985).

9. John Gardner & Timothy Macklem, *Reasons*, in THE OXFORD HANDBOOK OF JURISPRUDENCE AND PHILOSOPHY OF LAW 440, 474 (Jules Coleman & Scott Shapiro eds., 2002) (“[R]ationality . . . is simply the capacity and propensity to act (think, feel, etc.) only and always for undefeated reasons.”); see also JOSEPH RAZ, ENGAGING REASON: ON THE THEORY OF VALUE AND ACTION 1, 68 (1999) (“Being rational is being capable of acting intentionally, that is, for reasons . . .”; “An account of rationality is an account of the capacity to perceive reasons and to conform to them . . .”).

be contrary to preference when the time comes actually to do it. Consider the idea of a rational commitment. It might be rational for an agent to make such a commitment (*e.g.*, a threat, a promise) because the commitment makes alternatives available that are preferred by the agent to those alternatives that are available if no such commitment is made. Unfortunately, however, the same preference-maximizing rationality that is sufficient to motivate the idea of making the commitment initially is often also sufficient to undermine the commitment when it comes time to carry it out. For example, it might be wasteful or contrary to one's preferences, and nothing more, to carry out a threat (promise) if the threat (promise) has been unsuccessful (successful) in deterring (inducing) another's behavior in the way that was planned. To the extent that this *ex post* quandary is predictable *ex ante*, the threat (promise) is not a credible one to make, either for the party threatened (promised) or for the threatening (promising) party. The result is that the benefits of being able to make such threats (promises) are lost.

The problem is presented in a particularly stark way in the so-called "centipede game" which is represented in Figure 1.¹⁰ The bank has put out one hundred coins on a table. Two players, Art and Bart, are to take turns (at the nodes marked A and B respectively in Figure 1) removing either one or two coins from the table, each keeping all the coins that he removes. Removing one coin is represented as moving across to the right from one node to the next in the diagram; removing two coins is represented as moving down from a node. The game stops as soon as either player removes two coins, and at that point all the coins (and only those coins) still remaining on the table are returned to the bank. However, so long as each player takes only one coin (or moves across to the right), the game continues until all the coins are removed. Potentially, therefore, each player could take one coin at each turn and end up with fifty coins at the extreme right of the figure. Total payoffs (in the number of coins) to each of the two players at every possible end point of the game are represented by the numbers in parentheses, the first number being the payoff to Art, the second the payoff to Bart.

10. This game seems to originate with R.W. Rosenthal, *Games of Perfect Information, Predatory Pricing, and the Chain Store Paradox*, 25 J. ECON. THEORY 92 (1981). For recent discussion, see Robert J. Aumann, *On the Centipede Game*, 23 GAMES & ECON. BEHAVIOR 97 (1998); John Broome & Wlodek Rabinowicz, *Backwards Induction in the Centipede Game*, 59 ANALYSIS 237 (1999); and Wlodek Rabinowicz, *Grappling with the Centipede: Defence of Backward Induction for Bi-Terminating Games*, 14 ECON. & PHIL. 95 (1998).

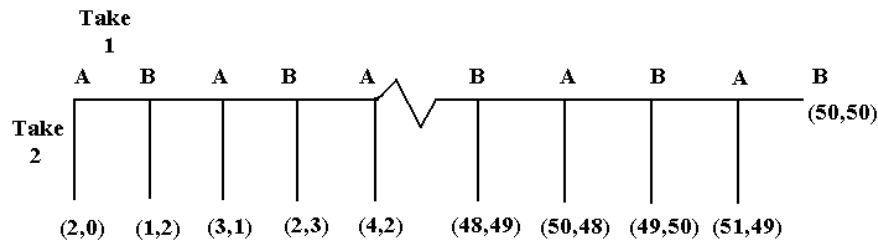


Figure 1: The Centipede Game

We are to imagine that Art and Bart are both rational in the sense that each wants to maximize his own monetary payoff from playing the game. Thus, each will not choose an option, or develop a strategy, if there is some other option or strategy that he could choose that will give him more money. Moreover, this rationality is common knowledge in the game. That is, each player knows that the other is rational, each knows that the other knows this, each knows that the other knows that he or she knows this, and so on.¹¹

Suppose Art has the first move (represented by the first A on the extreme left of the centipede in Figure 1). The rational choice theorist's standard argument, based on backwards induction, is that Art will rationally choose to take two coins (or move down in Figure 1) and the game will end. Of course, this seems a little problematic, even for Art; he might like to think that the game could have gone on a little longer so that he (and, incidentally, Bart too) could have picked up a few more of the one hundred coins that were available. But, unfortunately, that thought has no survival value under the assumptions of rationality and common knowledge of rationality.

To see why, imagine Art thinking ahead to where there are only two coins left on the table (that is, to the node marked A at the extreme right of the centipede in Figure 1). This means that up to this point in the game each player has taken only one coin and now has forty-nine in his possession. But now Art can either end the game by

11. The idea that some propositions might be common knowledge originates with DAVID K. LEWIS, *CONVENTION: A PHILOSOPHICAL STUDY* 52–60 (1969). For good discussion of common knowledge in general, and common knowledge of rationality in particular, see John Geanakoplos, *Common Knowledge*, in 2 *HANDBOOK OF GAME THEORY WITH ECONOMIC APPLICATIONS* 1437 (Robert J. Aumann & Sergiu Hart eds., 1994).

taking the two coins that remain (move down) or take only one (move across) and allow the game to end with Bart taking the only coin that is left. Clearly, the first option provides a higher total payoff for Art (fifty-one coins as opposed to only fifty) and, therefore, is the rational one for him. So that is the option he chooses on this play of the game and the game ends.

But now consider Bart thinking ahead to where there are three coins on the table,¹² that is, to the penultimate play in the game just before the one imagined by Art in the previous paragraph (or the last node marked B at the extreme right of the centipede in Figure 1). Since, under the assumption of common knowledge of rationality, Bart knows that Art is rational, he knows what Art will do in the next play of the game should Bart choose only one coin (move across) and the game move on to that next (and ultimate) stage. But Bart can do better than that by taking two coins (moving down) at this penultimate play (securing fifty coins rather than forty-nine), thereby stopping the game. So, being rational, that is what Bart chooses to do.¹³

Now, of course, this last choice by Bart is perfectly predictable by Art (again, given common knowledge of rationality) and so Art will anticipate at the pre-penultimate play of the game, when there are

12. Or, more accurately, consider Art thinking this about Bart. For all of this thinking is really going to an explanation of why Art, who has the first move in the game, will choose to take two coins on that first move. So it is really a question of what Art is thinking about what Bart is thinking (about what Art is thinking, etc.). All this is made possible by common knowledge of rationality.

13. There is, of course, a problem here that more than a few commentators have noticed. For the players to reach this point in the game, where there are only three coins remaining on the table, each player must have chosen not to terminate the game, that is, must have chosen to remove only one coin at all the prior turns. But, as the backwards induction argument goes on to show (under the assumptions of rationality and common knowledge of rationality, assumptions that the players themselves can use to generate the argument), to remove only one coin on one's turn is not rational. Thus, at the point when there are only three remaining coins on the table, for Bart to hypothesize that Art will remove two coins on his next move (should Bart take only one coin and allow the game to continue on to that next move) is for Bart to hypothesize that Art is rational on this next move even though, also by hypothesis, Art has shown no such rationality in the game so far. Is it plausible for Bart to have, or to hypothesize having, such a resilient (*i.e.*, contrary to fact) belief in Art's rationality? Indeed, is it plausible for Bart to anticipate acting rationally on his own turn, having himself acted irrationally in the game so far? More generally, is it plausible to argue or hypothesize, at *any* turn in the game, that the player (whose turn it is) will either act rationally at this turn, or believe the other player will act rationally on the next turn, if this turn could not have been reached except through irrational play either by himself or the other player (or both) at some point earlier in the game? For good discussion of this difficulty in the backwards induction argument, see Philip Pettit & Robert Sugden, *The Backward Induction Paradox*, 86 J. PHIL. 169 (1989). For a reconstruction of the argument that cleverly avoids this problem, at least at a formal level (by building in the assumption that all players believe that any turn in the game, *if it is actually reached*, must have been reached only by way of rational choices), see Broome & Rabinowicz, *supra* note 10. Also, for a similar argument, see Aumann, *supra* note 10.

four coins still on the table (that is, at the node marked A that is second from the right in Figure 1), that Bart will end the game at the next penultimate play. So, given that he is rational, Art will choose to do better by taking two coins rather than one at this pre-penultimate point, thus ending the game. And so on. We must conclude, therefore, that under this sort of inductive reasoning, and these assumptions, the game will end on the first play when Art takes two coins, leaving the other ninety-eight to be returned to the bank.

Does anything change if each player, at the beginning of the game, *promises* or commits to the other to only take one coin throughout the game? We can certainly see that each player has a reason for wishing that he could make such a sincere and credible promise (*i.e.*, a promise that the other player could rationally believe). After all, each would be so much better off if, by promising, each could induce the other to behave according to their respective promises; each would have fifty coins rather than Art having two and Bart having none. But the backwards induction argument, based on the assumptions that each player is rational and that this rationality is common knowledge, prevents the promise from being credible. Each player knows, under these assumptions and regardless of what has been promised by the other player, that it is rational for the other player to end the game on the next move should he himself choose, according to his promise, not to end the game by only taking one coin. Thus, it is pointless for each player to believe the other's promise, and just as pointless, therefore, for each player to make it.

Of course, there seems every reason to think that, in fact, two rational players in such a game would not actually play the game in the way that the backwards induction argument suggests. To accommodate this last point, the rational choice theorist's typical response has been to change the common knowledge of rationality assumption.¹⁴ That is, we will see the players play this game longer, and more profitably, the argument goes, because it cannot be assumed that each player knows that the other player is rational, or that each knows that the other knows that each is rational, etc. This change in the common knowledge of rationality assumption will allow Art (or Bart) at least to entertain the thought that at some point in the game he should take only one coin because the other player will not necessarily respond by taking two coins and ending the game in the next

14. See D. Kreps et al., *Rational Cooperation in the Finitely-Repeated Prisoner's Dilemma*, 27 J. ECON. THEORY 245 (1982).

round of play. Predicting exactly at what point the game might end depends on the precise details of how the common knowledge assumption is relaxed, and need not detain us here. The important point is that a relaxation of this assumption allows us to comprehend the thought that the players might play the game more profitably than they do under the strictest version of the backwards induction argument that is implied by assuming common knowledge of rationality.

Moreover, as an empirical matter, it does seem implausible to think that the players would actually have common knowledge of rationality. After all, such an assumption requires each player to know a great deal about the other player's rationality and, further, about the other player's knowledge about one's own rationality. Indeed, it requires a player to know about the other player's knowledge about one's own knowledge about that player's rationality! And so on. As the demands of common knowledge grow through these different levels, the assumption that there could actually be the sort of interpersonal transparency that is required seems more and more strained. And so it seems reasonable to the rational choice theorist to relax the common knowledge of rationality assumption.¹⁵

But I want now to suggest that the backwards induction argument does not depend so essentially on this sort of interpersonal knowledge. The argument, for all intents and purposes, will go through just as well if an agent is only required to have a sound knowledge of *his own* rationality and, in particular, if it is assumed that an agent knows that he cannot rationally intend or plan to do what he knows he will not rationally do (when the occasion arrives for him to act on that intention). To see this, consider the following variation on the centipede game.¹⁶ Suppose Perfectly Reliable Bart makes the following offer to Art: that at any point n in the game where it is

15. The economist seeks to relax the common knowledge assumption to explain the fact that players do not play the game in the way that the backwards induction argument suggests. Thus, while the argument might not apply as a contingent matter of fact, it is not as if they think there is anything problematic with the argument as such. Philosophers confronting the backwards induction argument are more inclined to think that there is something necessarily (not just contingently) wrong with the argument itself. Graham Priest has also noted that there is this difference in approach between philosophers and game theorists more generally. See Graham Priest, *The Logic of Backwards Inductions*, 16 *ECON. & PHIL.* 267, 268 (2000). For a good review of the broad range of philosophical arguments dealing with backwards induction, most of them dealing with so-called "surprise exam paradox," see ROY A. SORENSEN, *BLINDSPOTS* (1988).

16. This is a version of the variation introduced by SORENSEN, *supra* note 15, at 336–37. It builds on a problem about intentions introduced by Gregory Kavka in *The Toxin Puzzle*, 43 *ANALYSIS* 33–36 (1983).

Bart's turn he (Bart) will take only one coin so long as Art can form the intention *at n* to take only one coin on the next play of the game *n+1* when it is Art's turn. Bart is assumed here to be perfectly reliable in the sense that he always takes one coin at *n* on observing that Art has formed the requisite intention at *n*. In a sense, therefore, Bart is merely an automaton. Thus, there is no question here of Art having to make any difficult assumptions about Bart's rationality, let alone any higher level assumptions about Bart's knowledge of Art's rationality or, further, about Bart's knowledge of Art's knowledge of Bart's rationality. And, likewise, Bart does not need to know any of this about Art, although, for the purposes of the argument, Bart does need to be able to observe Art's intentions at any play of the game.¹⁷

Consider again the problem from Art's point of view. A new offer from Bart is only worthwhile to Art if he can form the requisite intention at that point to take only one coin on the next play of the game. But, at Art's last possible move in the game (that is, at the last node A on the centipede in Figure 1), when there are only two coins left on the table, Art knows he will take both of them. (After all, there is no possibility at this point of getting any new offers from Bart, and rational behavior, we assume, consists of maximizing one's monetary payoff). Thus, he knows, by assumption, that he cannot form the requisite intention at the move before this, when there are three coins on the table (and where it is Bart's move), to take only one coin on the next move. Thus, he knows that an offer from Bart at this point is worthless to him. But then, he asks himself, why not take two coins on the move (his second to last possible move, or the next to last A on the centipede) just before this move by Bart? Art would only not take two coins if, by taking one coin instead, he could again get Bart to make a worthwhile offer to him at the next move. But Art has already concluded that such an offer is worthless to him since he cannot form the requisite intention to make it worthwhile. So Art knows that it is pointless to take only one coin on his second to last move; he should take two. But then, of course, he cannot form the requisite intention at Bart's immediately prior move to take only one

17. Of course, there might appear to be something implausible about assuming that one person can observe another's state of mind, *e.g.*, another person's intentions. But not even this is really necessary. What is needed is only that Art *believes* this about Bart. However, as I hope now to suggest with this variation on the original example, the real implausibility of the backwards induction argument does not seem to turn on the particular version of interpersonal transparency that is used. The real problem appears to be in the notion of individual rationality that is being assumed.

coin. And so Bart's offer to him at that point is also worthless. But why then, he asks himself again, should he not take two coins at his third to last possible move? To take only one coin at this point only generates another worthless offer. In like manner it can be shown that all the prior offers that Bart might make to Art are worthless and that, as a consequence, Art will take two coins on the first move of the game. And none of this argument makes any general demands on Art's knowledge of Bart's rationality or vice versa. All that is required is that Art, being rational, know that he cannot form an intention to do something that he knows he will not rationally do.

This last requirement seems acceptable in general, but particular interpretations of it might not be. The real force of the requirement is in the idea that a rational agent cannot intend to do what he knows he will not do. But how does he know that he will not do it? Because, the argument goes, he knows that he is rational and that it will be *irrational* for him to do it. So far, so good; this much also seems acceptable. The difficulty arises on the interpretation of practical rationality that is used. If practical rationality means, simply, acting for (undefeated) reasons (and in rational choice theory this means acting according to reasons as manifested in all-things-considered preferences), then the requirement reduces to the idea that a rational agent cannot intend to do what he knows he has reason not to do. For then he knows he will not do it and this contradicts the real force of the requirement. But suppose that there was more to practical rationality than acting for reasons (or according to reasons as manifested in all-things-considered preferences). Then it would be possible for a rational agent to intend to do something that he had reason not to do. Why? Because *then* he might *not* know that he would not rationally do it even though he knew he had reason not to do it. After all, under this view, there is more to rationality than acting according to reasons. And without the knowledge that he might not rationally do it, he could intend to do it in a way that is consistent with the real force of the requirement.

Thus the possibility that there is more to practical rationality than acting for reasons opens up the further possibility that an agent can rationally intend to do what he has reason not to do. It is worth emphasizing that this is not the same as saying that he can have a *reason* to intend to do something that he has reason not to do. That was the possibility with which we began our investigation in this section of the paper. And we saw fairly quickly that an agent could have such

countervailing reasons; the examples of the centipede game and of promising seem to establish this point in a practically important way. What was problematic for the agent, however, was whether the reasons that he had for his prior intentions or promises could ever be made effective: could he actually form these intentions, or make these promises, if he had reason actually not to do as he intended or promised? The backwards induction argument, as applied either to interpersonal promises (as in the original centipede game) or intrapersonal intentions (as in the variation of the centipede game), suggested not. But now we can see that this argument turns on a questionable assumption that the rational choice theorist (amongst others) may have accepted too easily, namely, that practical rationality consists only in acting for reasons. For only then does the real force of the more general requirement, that an agent cannot rationally intend or plan to do what he knows he will not rationally do, reduce to the more particular idea that an agent cannot rationally intend or plan to do what he knows that he will have reason not to do.¹⁸

The suggestion here is that we should accept the real force of the general requirement, but not the particular interpretation of that idea that drives the backwards induction. This is because there *is* something more to practical rationality than acting for (undefeated) reasons (or acting for reasons as manifested in all-things-considered preferences). I hope that this section of the Article has given us some indication of why it might be important for an economic problem, the problem of rational commitment, that there is something more. The next section will tell us more specifically what that something more is.

18. In some very helpful comments on an earlier version of this argument, Wlodek Rabinowicz questioned whether it was plausible to impose this general requirement (*viz.*, that an agent cannot rationally intend or plan to do what he knows he will not rationally do). He suggested that even if an agent knew that he would not do *x* rationally when the time came actually to do it, the agent could nevertheless rationally intend or plan to do *x* if it was thought that forming the intention or plan would make it more likely that *x* would actually be done (albeit not rationally). It may even be that Ulysses binding himself to the mast to overcome (non-rationally) the lure of the sirens provides us with a classic example of such an effective and rational plan. However, in this sort of example, it seems that the physical restraint rather than the intention itself is doing the work to hold the agent to the plan. If Rabinowicz means to suggest that the mere fact of having adopted the intention or the plan, without more (such as using physical restraints, giving up hostages, etc., measures which either avoid the influence of reasons or change their balance at the moment of acting) can make it more likely that the act will be done, then he is closer to the structure of the problem being analyzed here. But then, as this Article will go on to argue in the next section, I am inclined to say that an act carried out under the normative requirements of an adopted intention or plan is rational rather than irrational.

III. REASONS AND THE NORMATIVE REQUIREMENTS OF PRACTICAL RATIONALITY

Theoretical reasoning, it is said, takes us from one belief state to another. Thus, if you begin by believing the proposition FG: “Frankfurt is in Germany,” and the proposition GE: “Germany is in Europe,” then theoretical reason would have you conclude by believing the proposition FE: “Frankfurt is in Europe.” Suppose that you do in fact believe FG and GE. Does this mean that you have a reason to believe FE? You may have reason to believe this (as it happens you do!), but not because of your beliefs about FG and GE. In fact, you might have no reason at all to believe GE or only have reasons not to believe GE. Thus, while it is true that *if* you believe FG and GE, you should *then* believe FE, there is nothing in this that gives you any *reason* to believe FE.

To see why, consider this alternative example. Suppose that you believe the proposition TG: “Toronto is in Germany” and the proposition GE: “Germany is in Europe.” Then, theoretical reason would have you conclude by believing proposition TE: “Toronto is in Europe.” But you have no reason, based on these beliefs, to believe TE. Indeed, you have many other reasons, independent of these beliefs, *not* to believe TE. And it is not that these other reasons, based on independent beliefs, simply prevail over, or outweigh, the reason you have to believe TE based on your beliefs in TG and GE. Rather, it is that there simply is no such reason to believe TE at all. *Any* independent reason not to believe TE would be enough to provide an all-things-considered reason not to believe TE, at least if the only “reason” that you claimed for believing TE was the fact that you believed TG and GE. This suggests that the weighing of conflicting reasons simply has no application here. The beliefs in TG and GE add nothing into the balance of reasons for believing TE.

But there does seem to be some sort of normative connection between believing TG (or FG) and GE and believing TE (or FE). What is that connection if it is not that believing the first two propositions provides a reason for why you should believe the third? John Broome provides an answer.¹⁹ Although your beliefs in the first two propositions provide no reason for you to believe the third, they do *normatively require* you to believe the third. Normative requirements differ from reasons, says Broome, in that they are *strict* and *relative*. They

19. Broome 1, *supra* note 7, at 401; Broome 2, *supra* note 7, at 105–06.

are strict because, in the context of theoretical reasoning, they really do *require* or obligate you to the conditional that *if* you believe TG and GE, *then* you should believe TE. If you believe TG and GE, but do not believe TE, then you are not entirely as you should be; in particular, you have failed to meet the normative requirements of rationality (here, the requirements of good theoretical reasoning). But these requirements, while strict, are relative because they do not detach from the conditional “if . . . then” and, therefore, do not give you any reason to believe TE *tout court*.

Reasons, on the other hand, are not relative in this way; they do detach and do give *independent* reasons, say, to believe TE (*e.g.*, perhaps a very reputable geographer told you that Toronto is in Europe). But these reasons are not strict; they are only *pro tanto*. That is, while you might have this independent reason to believe TE, it might still be that you do not believe it, perhaps because you have some other independent *stronger* reason for *not* believing it (*e.g.*, that TE goes against everything you were taught in school). However, because reasons are not strict, not believing what you have a reason to believe is quite consistent with being entirely as you ought to be. While there might be *a* reason to believe TE, the *balance* of independent or detached *pro tanto* reasons might be such that you do *not* believe TE. But that is no problem.

Reasons, therefore, are weaker than normative requirements in being only *pro tanto* and not being strict. But they are stronger than normative requirements in being independent rather than relative. These are differences that go to the very logical structure of each. We are now ready to see how these important logical differences are relevant to practical reasoning and what they can do for an agent.

Practical reasoning differs from theoretical reasoning in that it concludes in a state of mind that involves a decision or intention (usually, to act) rather than a belief.²⁰ Here is an example:

- (1) I intend that [I will visit Heidelberg];
- and (2) I believe that [To visit Heidelberg I need to fly to Germany];
- and so (3) I intend that [I will fly to Germany].

The bracketed propositions provide the content for the different statements and the prior non-bracketed terms reveal my state of mind, or attitude, with respect to each of the propositions. The logic

20. Broome 3, *supra* note 7, at 175.

of the reasoning is contained in the propositions themselves.²¹ We can see this if we think of these same three propositions under the aspect of theoretical reasoning, where only belief states of mind apply. If I believe the bracketed proposition in (1), and I believe the bracketed proposition in (2), then the ('and so') logic of theoretical reasoning will have me conclude that I believe the bracketed proposition in (3). In the practical reasoning that is described by the above example, the same ('and so') logic applies, although now it takes us from a prior intention state of mind in (1) and the belief state of mind in (2) to the concluding or derivative intention state of mind in (3).

We can now pose questions about practical reasoning that are fully analogous to the ones that we posed earlier about theoretical reasoning. Does my prior intention in (1) together with my belief in (2) give me any *reason* for my final derivative intention in (3)? No, not any more than the same logic applied to the following three statements would give you any analogous reason to have the derivative intention in (6):

- (4) I intend that [I will visit Heidelberg];
- and (5) I believe that [To visit Heidelberg I need to fly to Canada];
- and so (6) I intend that [I will fly to Canada].

The prior intention in (4) together with the belief in (5) gives me no reason to have the derivative intention in (6).

But it is true that I am *normatively required* to have the intention in (3) (or in (6)) *if* I have the intention in (1) (or in (4)) and the belief in (2) (or in (5)). While relative in this way, this normative requirement of practical rationality is, as all such normative requirements are, strict. In other words, if I do have the intention in (1) (or in (4)) and the belief in (2) (or in (5)), then, if I do not have the intention in (3) (or in (6)), I am not entirely as I should be. In particular, I have failed to meet the normative requirements of practical rationality.

These are, by now, familiar enough points. So let us add a little conflict into the mix. Suppose that I do have the intention in (1) and the belief in (2). Then I am normatively required to have the intention in (3). If I don't, I am not entirely as I should be. But suppose that I have an independent reason *not* to have the intention in (3) and, further, no independent reason to have it. (Perhaps there is a strike by air traffic controllers in Germany, making any flight to

21. Broome 4, *supra* note 7, at 89.

Germany less safe.) Then the strict normative requirements of practical rationality are in conflict with my independent *pro tanto* reasons. Am I still entirely as I should be? It seems not. Something is wrong here and needs sorting out.

Here is where the *relative* quality of normative requirements of practical rationality can be useful. The strict quality of these normative requirements obligates me to have the derivative intention in (3), but only *if* I have the prior intention in (1) and the belief in (2). Thus, I can satisfy these strict requirements either by accepting the antecedent conditions of the conditional *and* accepting the consequent (*modus ponens*), or by rejecting the consequent *and* rejecting one or other (or both) of the antecedent conditions that require the consequent (*modus tollens*). The fact that I have an independent reason for rejecting the consequent seems to provide me with some motivation for the second method of satisfying the normative requirements of practical rationality. Then I could satisfy both my independent reason for not having the intention in (3) *and* the strict normative requirements of practical rationality. And, after this adjustment, I would be entirely as I should be.

Suppose, as seems reasonable, I cannot adjust my beliefs in (2).²² Then to make the necessary adjustment I would need to change or repudiate my prior intention in (1).²³ But that does not seem problematic, at least on the argument so far. So far I have not provided any reason for my prior intention in (1); there is only the fact that I have it. But it seems implausible that the mere fact of having this prior intention could count for much if I have an independent reason not to have the derivative intention in (3). This is consistent with the insight that a prior intention in (1), together with the belief in (2), gives me no reason to have the derivative intention in (3). Thus, while the normative requirements of practical rationality strictly require me to have the intention in (3) if, as a matter of fact, I have the intention in (1), they do not provide much normative resistance against my changing that fact by repudiating the intention in (1).

22. On the difficulty of deciding to believe, see BERNARD WILLIAMS, *Deciding to Believe*, in *PROBLEMS OF THE SELF: PHILOSOPHICAL PAPERS 1956-1972*, at 136 (1973).

23. Broome 2, *supra* note 7, at 112. Note that for Broome, repudiation is more than merely ceasing to have the prior intention, but it might not require a reason either. Suppose there was no reason for the prior intention. Why, then, should it take a reason to give it up? Broome requires repudiation to be deliberative, but not necessarily "with reason," something that is a little mysterious.

But now suppose that there were reasons for adopting the prior intention in (1), and not merely the fact that I had adopted it. This is more analogous to the situation in my variation on the centipede game where Art's adopting the prior intention to take only one coin (supposing he could form such an intention) seems to be quite an advantage for him. Now it is less clear that either Art or I should simply repudiate the prior intention once an independent reason not to have the derivative intention in (3) shows itself. For now it seems that there are reasons pulling in opposite directions, namely the reason that supports forming the prior intention and the reason that has now surfaced for repudiating it.

It is tempting to think that these conflicting reasons must somehow be balanced or weighed against one another. However, if this was the only possible approach to the problem, the result of all this discussion about the logical (categorical, qualitative) difference between reasons and normative requirements would be disappointing indeed. For then, in any serious case of conflict between reasons and normative requirements, rationality would seem, after all, to reduce to acting on a mere quantitative balancing of *pro tanto* reasons. Further, in the face of such a balancing exercise one would surely have to face the question that the backwards induction argument presents so forcefully: what weight should be given to the reasons we once had for adopting the prior intention against the "reasons" we now seem to have for repudiating it? (Strictly, the latter are not really reasons, of course, but only the rational motivations that arise out of a concern for satisfying the normative requirements of practical rationality.) Should not our current "reasons," that is, the rational motivations that still have some application or force, simply prevail (without regard to weight) over reasons that are now defunct?

However, I want to suggest that a more categorical or formal (less quantitative) approach to the problem is possible. Consider two different situations:

- (a) where the reason for repudiation was anticipated and accounted for when the prior intention was adopted; and
- (b) where the reason for repudiation was not so anticipated and accounted for.

Under situation (b) it would seem to be irrational, indeed, almost thoughtlessly mechanical to go ahead and act (or, more accurately, form the derivative intention to so act) on a prior intention without allowing these new unanticipated (and truly independent) reasons to

have any impact. After all, it is a mistake to think that the whole of practical rationality is action according to normative requirements. There is also rational repudiation in the face of independent reasons.

But, equally, under (a) it would seem to be irrational, in the face of (now countervailing) reasons that had already been anticipated, considered, and accounted for under the adoption of the prior intention, not to carry out that prior intention in one's derivative intention. After all, it is also a mistake (the one identified by Broome) to think that the whole of practical rationality is action according to reasons. And once these reasons have already been accounted for, there seems only to be the normative requirements of practical rationality to be considered. This is the stuff of rational commitment.

Thus, if there is more to practical rationality than acting in accordance with the (balance of) reasons, and in particular if there is the practical rationality of acting under normative requirements as well, then there is the rational possibility of forming an intention to do what one knows (precisely because one has anticipated the reason) one will have reason not to do. Therefore, more specifically, there is the rational possibility of forming the intention to carry out a promise or threat, or to take one coin in my variation of the centipede game, even if one knows, when it comes time to carry out this intention, that it will be contrary to one's preferences (or reasons) to do so. The important point is that these countervailing preferences will already have been anticipated and accounted for when the prior intention was formed. Further, having formed the prior intention to carry out a promise or threat, and without any independent (*i.e.*, unanticipated) reasons for repudiating it, it only remains to meet the normative requirement of practical rationality in actually carrying out this intention in one's actions. This is, of course, what the alternative account of the rational actor promised to provide. The next section will show that this alternative account is essentially law's account of the rational actor.

IV. LAW'S RATIONAL ACTOR: ADJUDICATING DEFEASIBLE RULES

A system of common law is more than a mere list of all the decisions that judges have chosen to impose upon litigants. It is also comprised of the legal rules which are said to bring order to these different results. Of course, within the common law method of adjudication, these rules do not typically appear in some pre-existing authoritative text, like the rules of a tax code. Rather, they develop over

time in the cases, sometimes abruptly, more often gradually, as the general rules of, say, tort or contract law. This might suggest that the general rules are mere descriptions of the behavioral regularities of judges. After all, if they do not pre-exist the cases, then the only other option seems to be that the rules come into place as rationalizations for what has actually been done for some independent reason in the particular case.

Now it is certainly true that, like the laws of science, rules of law bring order and understanding to a legal reality which is independently laid down and which can be the object of external and scientific observation, the stuff of induction. But, more than this, rules of law are also said to bring order to a judge's self-conscious understanding of what she does and, further, of what she feels she ought to do. Rules, it is said, help to provide particular justifications for the legal result that she comes to in a case. Rules, therefore, can be said to order the law from both an external *ex post* point of view (the point of view of the scientific observer) and from an internal *ex ante* point of view (the point of view of a committed participant or judge guided by rules in the legal decisions she makes).²⁴

However, that legal rules have this double aspect suggests that there will be some ambiguity as to what the proper relationship should be between these rules of law and their apparent instantiations in the cases. Under the more descriptive, scientific account, the particular case sets the standard for the rule. A rule will fail as a description, or fail to provide a proper understanding, in so far as it is an inaccurate representation of what is going on in the case. Although a limited number of exceptions can sometimes be said to "prove" the rule (since that is what the very idea of an exception presupposes), too many will be fatal to its descriptive claim.

On the other hand, under the more prescriptive account, where rules are said to provide reasons or justifications for judges to decide cases one way rather than another, the relationship between a case and the applicable rule is reversed. Now the rule sets the standard for the case. Moreover, because the rule has this seeming autonomy from the cases, it can pronounce almost any number of them as "wrongly decided." The number of such decisions only attests to the frequency of judicial error, leaving the legal rule intact and still perfectly capable of governing other cases.

24. See H. L. A. HART, *THE CONCEPT OF LAW* 55 (1961), for a discussion of the importance of this "internal point of view" for the committed participant in law.

To accommodate this dual aspect of common law rules, what is required is an account that allows the rules to be strong enough to guide judicial decision-making in particular cases, but not so strong that it does not allow for the possibility that these same rules might require revision in the light of these same cases. This might suggest that what we are looking for is something quite banal, namely, an account of rules that merely *weighs* the good that they do as rules (say, in securing general expectations, making life more predictable, controlling harmful judicial discretion, etc.) against the independent good that can be done by revising or abandoning the rule in some particular case. But I want to press the intuitive point that a rule, and the following of a rule, is more strict (or more “rule-like”) than this balancing or weighing metaphor allows. Borrowing from the above analysis, I want now to argue that a legal rule can *normatively require* a particular result in some case without regard to the good that is achieved in, or frustrated by, that result, so long as this good (and its possible frustration) has already been anticipated by the rule and is accounted for in it. But a good that is frustrated by that result can also be an *independent reason*, even a decisive independent reason, for *not* following the normative requirements of a rule. A reason would be independent in the required sense if it involved a good that was not anticipated by the rule, and accounted for in it. Thus, an integration of the normative requirements of practical rationality with the idea of independent reasons promises to provide the right combination of rule-respect and rule-denial that is required if we are properly to accommodate the essential roles that rules and cases each play within a theory of common law decision-making.

This way of integrating the normative requirements of practical rationality with independent reasons is familiar enough to those who understand the common law as a system of defeasible rules. Mention of defeasibility, of course, reminds us of H.L.A. Hart, as Hart was influential in introducing the idea of defeasibility into legal theory.²⁵ Borrowing the idea from the law of property, Hart noted that “a legal interest . . . is subject to termination or ‘defeat’ in a number of differ-

25. H. L. A. Hart, *The Ascription of Responsibilities and Rights*, in *ESSAYS ON LOGIC AND LANGUAGE* 145–66 (A. Flew ed., 1960). There is now a large literature arguing that the formal structure of legal reasoning is to be found in defeasible (or non-monotonic) logic. See, e.g., HAGE, *supra* note 5, and the references cited there. See also Bruce Chapman, *Law Games: Defeasible Rules and Revisable Rationality*, 17 *LAW & PHIL.* 443, 446–447 (1998); Bruce Chapman, *Law, Incommensurability, and Conceptually Sequenced Argument*, 146 *U. PA. L. REV.* 1487, 1507–26 (1998); Frederick Schauer, *On the Supposed Defeasibility of Legal Rules*, 51 *CURRENT LEGAL PROBS.* 223 (1998).

ent contingencies but remains intact if no such contingencies mature.”²⁶ Although he believed that this idea, or more particularly the dual structure of this idea, had wide application in the law, Hart developed the idea most explicitly with reference to the concept of a contract. He might equally have referred to a rule of contract formation. For Hart, as much as for other legal scholars, there is the usual list of positive conditions required for the existence of a valid contract (*e.g.*, at least two parties, an offer by one, its acceptance by the other, consideration on both sides). However, knowledge of these conditions does not, according to Hart, give us a full understanding of the concept of contract nor of the rule of contract formation. We also need to know the various ways in which the claim that there is a contract (under the concept or the rule) can be defeated. Such defenses to the claim might include, for example, that there was fraudulent misrepresentation, duress, or lunacy. Hart argued, therefore, that the concept or rule of contract formation was best captured structurally as a list of conditions that are *normally* necessary and sufficient for the existence of a valid contract, together with a series of “unless” clauses that spell out the conditions under which this existence claim is defeated. However, Hart recognized that the list of “unless” clauses could not, in all likelihood, be exhaustively specified. And so such a rule would often end only (and perhaps only implicitly) with the word “unless. . . .” However, Hart was clear: “A rule that ends with the word ‘unless . . .’ is still a rule.”²⁷ Specifically, it is a defeasible legal rule.

Now the idea that a rule qualified by an “unless” clause is some special sort of rule has met with a good deal of skepticism. Moreover, the attack appears to come from two different directions. On the one hand, the claim is that once the “unless” clause is written (even implicitly) into the rule, we simply have a more complicated rule.²⁸ It might be a rule with more limited scope, but it is a rule nonetheless. Defeasibility plays no essential role. On the other hand, if the “unless” clause simply signals the occasion to modify the rule in light of some background justifications (or reasons) for the rule, then the background justifications are all that really apply and the force of the

26. Hart, *supra* note 25, at 148.

27. HART, *supra* note 24, at 136.

28. RONALD DWORKIN, TAKING RIGHTS SERIOUSLY 24–25 (1977) (accounting for exceptions to a rule under an articulation of a more complicated version of the rule).

rule *qua* rule disappears.²⁹ Thus, a defeasible legal rule is reduced either to a rule without defeasibility or to defeasibility without a rule.

However, these attacks miss their mark. To see why, suppose that there is some general rule and a situation arises which, while not precisely anticipated, clearly lies within *both* the linguistic contours of the rule *and* its background justification. Then it is tempting to say that the rule will simply be applied, for the question of its defeat in the face of a countervailing background consideration does not arise. But the critic of Hart's notion of defeasible legal rules, and more particularly that critic who seeks to reduce all rules to their background justifications, will think that this is not a very interesting case. The rule is doing no real normative work here, she will argue; everything is being done by a direct application of the justification or purpose that the rule serves and which, in this sort of situation, is in agreement with what the rule requires.

But this is a mistake. The rule does do its own normative work here, and it does so under the aspect of a normative requirement of practical rationality and not under the direct application of the background justification or independent reason for the rule. The promising example, discussed earlier, makes this clear. Suppose we had a rule for seeing to it that our promises to others were performed as intended.³⁰ The reason or background justification for the rule might well be the welfarist one that we are better off performing our promise as intended than we would be if we made no such promise at all, even if (we anticipate that) actually carrying out that promise makes us all worse off than we would be if we secured the benefits of making the promise and did not incur the costs of actually performing it. This presents us, of course, with a familiar problem. We have already seen that we might have a particular background justification or reason, which we can call reason *W* (for welfare), for intending to carry out a promise, or (now) having a promising rule, which can *also* furnish us with a reason *not* to follow the rule when the particular situation arises for carrying out the promise as previously intended. Nevertheless, as a matter of normative requirement, we should follow, or 'simply apply', the rule because this is precisely the situation that was anticipated by the rule and was already accounted for in it. Moreover, we properly (rationally) follow the rule in the particular case even

29. Schauer, *supra* note 25, at 232–33.

30. The promising example is discussed both by HART, *supra* note 24, at 136; and by Schauer, *supra* note 25, at 224, 229.

though there is (now) some cost in terms of reason *W* (which provides the background justification for the rule) in doing so. Thus, it is wrong to think that at the point where the rule is to be followed we are simply appealing directly to the background justification. On that view, a view that (here) reduces practical rationality to acting in accordance with reason *W*, we would *not* follow the rule. But we do follow the rule, and we follow it as a matter of normative requirement. Further, we have no *independent* reason for not following the rule, having already anticipated and accounted for the countervailing consideration in terms of *W* *within the rule itself*.

With the practical significance of a rule secured in the case where the rule and the background justification appear to agree, we have only to recognize that in the alternative sort of situation, where some event occurs that appears to lie within the rule's linguistic contours, but which was not anticipated under the reasons that provide for the rule's background justification, there *is* an independent *pro tanto* reason for not following the rule. After all, under this more purposeful (content-sensitive, reason-based) understanding of the rule, the rule simply does not apply and there is no normative requirement actually to follow it. Indeed, there is only an independent reason for not following it, and for re-formulating the rule (perhaps with a further "unless" clause) in light of the new (unanticipated) *pro tanto* reason (or, perhaps, with a view to some new balance between new and old *pro tanto* reasons). Thus, the combined application of normative requirements and independent reasons makes sense of Hart's claim that both rules *qua* rules and the defeasibility of rules can sensibly be integrated into a full account of practical legal rationality. And it also allows us to comprehend the idea, alluded to at the beginning of this section, that particular cases can, apparently simultaneously, both determine legal rules (as a matter of independent reason) and be determined by them (as a matter of normative requirement).

A cautionary note is in order before concluding this section. While this analysis provides for reasons and rules to have formally (or logically) separate roles within law's account of the rational actor, a separation that we have argued would be useful to the economist for resisting the force of the backwards induction argument, as a practical matter it will be difficult to separate reasons from rules. This is because, on the analysis provided here, the reasons, while substantive, will be rule-based, and the rules, while formal, will be reason-based. Consider each of these claims in turn.

(a) *Rule-based reasons*: Judges, when they concern themselves with reasons *qua* reasons in particular cases, will concern themselves less with the substantive reason behind some prior commitment or rule, and more with the (independent, detached) reasons that might arise for narrowing or broadening the prior commitment or rule in light of developing circumstances. In other words, the judge, even when she is working with substantive reasons, will be working on the margin of some more general rule adopted for application in the cases, pressing forward for its application in the case *unless* she can be convinced by litigants (who will be addressing these margins) that the case is unlike those already anticipated under the rule. In this respect, the judge's use of substantive (independent) reasons is *rule-based*, although not *rule-bound*. In short, there is a rule, but it is defeasible.

(b) *Reason-based rule*: Judges, even when attempting to apply some prior rule or commitment *as a rule or commitment* in a given case (that is, as a matter of formal normative requirement rather than substantive independent reason), will often have to attend to the reasons that motivated the adoption of the rule or commitment in the first place. For without some attention to these reasons, they run the risk of not having a full sense of what the rule really is and what it requires. This is the stuff of a purposeful, or *reason-based*, interpretation of the rule. But the purposeful application of the rule is still formal, a matter of strict normative requirement, and not substantive. It is to be distinguished from applying the reasons for the rule directly to the facts of the case.

V. CONCLUDING REMARKS

In the introduction to this Article I committed to providing a “legal analysis of economics” as a counterbalance to the “economic analyses of law” which are so much more prevalent in law and economics scholarship. I like to think that I have followed through on this commitment in a way that is consistent with the strict normative requirements of practical rationality. Part II of my Article provided an economic subject matter for analysis, the problem of rational commitment. Part III suggested an alternative account of the rational actor as someone who acted under a richer conception of rational conduct, a conception that embraced rational action as action according to (undefeated) reasons, or reasons as manifested in (all-things-considered) preferences, *and* action that respected the normative requirements of practical rationality. I argued that this alternative

conception would allow an actor to follow through on commitments rationally made even if there were substantive reasons (so long as they were *anticipated* reasons) to defect from these commitments. Finally, in part IV I argued that this richer conception of practical rationality was manifested in the common law adjudication of defeasible legal rules, where rules, apparently simultaneously, both determine particular cases (as a matter of strict or formal normative requirement) and are determined by them (as a matter of independent, substantive reason).

However, as I pause at the end of the argument that I like to think fulfills the promise that I made, I have to wonder whether this is not also a moment to revise the commitment (at the margin) so as to incorporate an unexpected event. As I indicated, I set out to provide a legal analysis of an economic problem with the thought that the analysis would continue to be autonomous of its subject matter in the same way that economic analysis is so often unaffected by the legal subject matter that it engages. There would be some symmetry (and, admittedly, some retributive satisfaction against the more standard economic analyses of law) in such a legal analysis of economics. But I have to admit that a consideration of the economic problem of rational commitment, even under the lens of a legal analysis, has given me a better understanding of the precise structure of defeasible legal reasoning than I had before. So a somewhat revised articulation of my commitment, or of what in the end I seem to have provided under my commitment, is needed: it is not so much that I have provided a purely *legal* analysis of a purely *economic* problem as a *law and economics* analysis of a *law and economics* problem. But that sounds to be exactly what an equal partnership of law and economics should provide, not only because it manifests more equality between the disciplines and draws on a richer range of analyses, but also because it provides for a closer fit between the analysis and its subject matter. Anything more distant risks misunderstanding the latter under the former.