Norms, Repeated Games, and the Role of Law

Paul G. Mahoney† & Chris William Sanchirico‡

TABLE OF CONTENTS

Norms, Repeated Games, and the Role of Law

Paul G. Mahoney & Chris William Sanchirico

One of the major developments in legal scholarship over the last decade has been a shift of attention away from formal legal rules toward informal, decentralized methods of social control, or social norms. Many scholars suggest that social norms, not legal rules, are the mainstay of social control. Such a view requires a theory of why individuals would follow norms against their immediate self-interest without threat of formal legal sanction. In seeking an explanation, the norms literature draws heavily on the game theoretic idea that individuals follow norms because of the possibility of community retaliation. Norms scholars express concern, however, that such threats are not credible because there is a free rider problem in inducing community members to engage in costly enforcement. We demonstrate that this "third-party enforcement problem" is illusory. The norm "tit-for-tat," which has received almost exclusive attention in law and norms literature to date, is indeed plagued by this problem. But many other simple and plausible norms are not. We define a norm called "defect-for-deviate," or "def-for-dev," which is similar in spirit to tit-for-tat, but is structured so that the punishment of norm violators is always in the individual interest of the punisher.

Yet there are other important reasons for skepticism about game theoretic approaches to social control that norms scholars have not recognized. We highlight the "counterfactual problem": the fact that the game theory of norm enforcement requires individuals to continue to believe that their community has adopted the norm, even in the face of proof that this belief is false. The counterfactual problem opens up avenues for law that the literature has not yet identified. We contend that law does not simply help players arrive at a normative equilibrium, but is required to sustain that equilibrium. This observation has the virtue of consistency with actual patterns of law enforcement.

INTRODUCTION

Over the past decade, legal scholars have devoted considerable attention to informal, decentralized methods of social control, or social norms. A large portion of that scholarship references the theory of repeated games,

typified by the repeated prisoner's dilemma.[1] This makes sense, as repeated game theory is concerned with the possibility of implicit, "self-enforcing" agreements. Economic theories of social norms, in the abstract, are concerned with the same thing.

Norms scholars, however, have been reluctant to adopt wholesale the lessons of repeated game theory. In part, this reluctance stems from the perceived reductionism and empirical paucity of the pure game theoretic approach. But norms scholarship also expresses objections to repeated game theory on its own terms. A prominent example is the so-called "third-party enforcement problem." The third-party enforcement problem arises because inducing individuals to cooperate against their short-term interest requires the threat of future punishment. When cooperation is enforced by norm rather than law, those punishments are carried out by other members of the community via, for example, ostracism. Norms scholars question whether community members will really carry out their implicit threat to punish those who are meant to be ostracized. Punishment is often costly for the punisher, as when profit can be made from dealing on the sly with the deviators. Accordingly, inducing punishment seems at first glance as difficult as inducing cooperation in the first place, leading to a "second-order" prisoner's dilemma.

However, as game theorists are aware, there is no third-party enforcement problem—at least not as characterized in the norms literature to date. The problem, as conventionally understood, disappears if we revise our repeated game solution concept from simple Nash equilibria to Nash equilibria that are also subgame perfect.[2] Put simply, subgame perfection implies that all threats are credible because it is in the player's best interest to carry them out, even if doing so is costly in the short run. Most important for current purposes, the *N*-player repeated prisoner's dilemma has many subgame perfect equilibria that rely on third-party enforcement to induce cooperative play.[3]

---

1. In a repeated prisoner's dilemma game, players meet for a pre-specified number of rounds and can condition current actions on the past history of play. For a fuller description of the repeated prisoner's dilemma and its equilibria, see *infra* Appendix: The Repeated Prisoner's Dilemma.

2. *See* Reinhard Selten, *Spieltheoretische Behandlung eines Oligopolmodells mit Nachfrageträgheit* (pts. 1 & 2), 121 ZEITSCHRIFT FÜR DIE GESAMTE STAATSWISSENSCHAFT 301, 667 (1965); R. Selten, *Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games*, 4 INT'L J. GAME THEORY 25 (1975). Jonathan Bendor and Dilip Mookherjee make a similar point about the treatment of the third-party enforcement problem in the political science literature, though their argument also seems to have gone unheeded in the law and norms literature. *See* Jonathan Bendor & Dilip Mookherjee, *Norms, Third-Party Sanctions, and Cooperation*, 6 J. L. ECON. & ORG. 33 (1990).

The concepts of Nash equilibrium and subgame perfection are discussed more fully *infra* in Appendix: The Repeated Prisoner's Dilemma.

3. Like the law and norms literature, we focus on the prisoner's dilemma in this Article. Qualitatively, however, our analysis applies to any repeated game with a "bad" equilibrium and a mutually beneficial outcome that may not be incentive-compatible in a one-shot context. Indeed, the

The prior literature's concerns about the third-party enforcement problem likely stem from its focus on a particular repeated game strategy—the well-known "tit-for-tat" strategy.[4] Tit-for-tat, as we establish, is not subgame perfect for any discount factor. As a consequence, it relies on third-party (and second-party) threats that are not credible.[5]

Our first objective in this paper, therefore, is to suggest an alternative canonical repeated game strategy. In defining this strategy, we distinguish "defection" from "deviation." Defection and cooperation are the two action choices in any single repetition of the prisoner's dilemma. Deviation, by contrast, is defined with respect to a given strategy for the repeated game. A repeated game strategy tells the player what to do in any given round based on the history of play up to that point. To deviate from that strategy in any given round is to fail to follow its instruction. Thus, a defection is not a deviation if the repeated game strategy calls for defection in the current round.

Tit-for-tat punishes defection in the prior round with defection in the current round. As typically defined, however, tit-for-tat does not distinguish between provoked and unprovoked defections. Thus, in a two-player game, if Player A defects in period five, Player B is instructed to defect in period six. When Player B defects in period six, however, the strategy instructs Player A to defect in period seven. The game thus settles into an "echo" in which the players take turns defecting. Equally troubling, tit-for-tat does not call for punishment of unwarranted cooperation. Thus, in an

---

existence of subgame perfect equilibria supporting cooperative payoffs in the N-player repeated prisoner's dilemma generalizes to the existence of subgame perfect equilibria supporting virtually any payoff vector in the one-shot game. For details, see DREW FUDENBERG & JEAN TIROLE, GAME THEORY (1991).

We focus on the prisoner's dilemma for the same reason that the law and norms literature focuses on it. The prisoner's dilemma is a simple, useful, and familiar representation of situations in which mutual cooperation is mutually beneficial, but not individually incentive-compatible. This conflict between social welfare and individual rationality maps well onto many situations that involve externalities. Among these are some of the basic underlying games of law. Respecting property rights, honoring mutually beneficial promises, and accounting for the safety of others can each be fruitfully analyzed as a prisoner's dilemma. To the extent that these situations do not map onto the particular configuration of payoffs in the prisoner's dilemma, our analysis will apply as well to whatever payoff structure provides a more accurate description.

4. *See, e.g.*, Jonathan R. Macey, *Public and Private Ordering and the Production of Legitimate and Illegitimate Legal Rules*, 82 CORNELL L. REV. 1123, 1130 (1997); Richard H. McAdams, *Cooperation and Conflict: The Economics of Group Status Production and Race Discrimination*, 108 HARV. L. REV. 1003, 1012 (1995). For a description of tit-for-tat, see *infra* Appendix: The Repeated Prisoner's Dilemma, pt. B.

5. The fact that tit-for-tat is not subgame perfect has received little, if any, attention in the law and norms literature. However, the point is known within game theory. For example, Eric Rasmusen raises the issue in the context of two-player repeated games without time discounting. ERIC RASMUSEN, GAMES AND INFORMATION 120 (1989). Rasmusen, however, stops short of proving the point. The principal drawback of his explanation is that it does not discount payoffs, but provides no alternative means of differentially ranking infinite sequences of positive payoffs, all of which sum to the same infinity. *See id.*

N-player game, if Player D is supposed to defeet in the current period (in order to punish a prior-period defector) but chooses to cooperate instead, tit-for-tat does not instruct the remaining players to punish D in the next round for failing to be a "team player."

Our alternative strategy is designed to avoid these drawbacks by punishing deviation from expected behavior rather than defection. It does not, accordingly, punish defection when that was the strategy's instruetion for the prior period. It does, however, punish cooperation in that situation. We call the strategy "def-for-dev" to emphasize this distinction. Def-for-dev, we argue, is a more reasonable metaphor for the real-world behaviors that norms scholars analyze than is tit-for-tat. Moreover, unlike tit-for-tat, def-for-dev is subgame perfect for a range of discount factors.

Although the third-party enforcement problem is illusory, there are still reasons for skepticism about the application of mainstream repeated game theory to law and norms. Our second objective in this paper is to draw attention to a problem that receives less attention in the game theory literature but is a more serious challenge to the practical application of that theory to actual systems of social control.

The problem concerns the self-contradictory nature of self-enforcement. In order for a norm to be self-enforcing, it must prescribe punishments that are adequate to deter violations and it must also induce community members to mete out those punishments.[6] When punishing is costly, inducing community members to punish requires that they believe that if they fail to do so, others will punish them for failing to do their duty. The souree of this belief is the would-be punisher's confidence that the norm remains intact in the wake of the violation that he is now instructed to punish. But the violation itself ought to shake this confidence. Had the original violator believed the norm to be in effect, she would not have violated it. After all, the norm prescribes punishments that are adequate to deter violations. The original violation, therefore, may cause the would-be punisher to question whether the norm remains intact. It may also cause him to question whether his fellow community members—who enforce his current duty to punish—believe the norm to be intact. His doubts about the continued viability of the norm may lead him to conclude that he, too, can shirk his duty by failing to punish. Indeed, this instability of the norm in the face of deviations can become a self-fulfilling prophecy. Imagining all this ahead of time, the original norm violator might have calculated that her

---

6.   When the literature models social norms as repeated game equilibria, the focus is on social norms as "informal social regularities that individuals feel obligated to follow . . . because of a fear of external non-legal sanctions . . . ." Richard H. McAdams, *The Origin, Development, and Regulation of Norms*, 96 MICH. L. REV. 338, 340 (1997). Because we are speeifically interested in evaluating the literature's use of repeated game theory, we adopt this definition of the term "norm." For our purposes, nothing is gained by a more general or detailed definition of this hotly eontested term. For a lucid discussion of various possible definitions, see *id.* at 350-51.

deviation could cause others to abandon the norm and therefore decline to punish her.

While the theory of games has devoted much attention to the conventional third-party enforcement problem—insuring that punishment is in the interest of the punisher given common belief that the norm (or strategy profile) is still being played—it has essentially ignored the possibility that deviations may well unsettle such common beliefs. The theory of games typically assumes without discussion that players remain certain that an equilibrium strategy profile is being played even after events occur that call this fact into question. This contradiction, which we call the "counterfactual problem," challenges the theoretical structure upon which norm scholars have rested their theory of self-enforcing norms. The counterfactual problem, and not the conventional third-party enforcement problem, is a legitimate locus for concern regarding game theory's applicability to real-world interaction.[7]

At first glance, our observations about what is and is not a problem may seem like esoterica, but we raise them because they have, in turn, a consequence that is central to the purpose of legal scholarship. Legal scholars write about norms in order to illuminate the relationship between norms and law. The puzzle that has motivated the development of the field is that law, by itself, appears to do too little. Legal sanctions in the real world are too light and too infrequent to guarantee that harmful, self-interested behavior will never pay. Something else, then, must be responsible for the fact that most people are reasonably cooperative most of the time. That something, in the view of many scholars, is the pressure of community norms enforced by informal, decentralized sanctions.

But if norms can induce cooperative behavior, what role is left for law? If all or most social dilemmas can be solved without an external enforcer, then the puzzle is not why there is so little law, but why there is so much. Some of the norms scholarship, therefore, tries to recreate a role for law. Most commonly, law is thought to act as a focal point that helps players to find the correct strategy or a catalyst to help them jump from an inefficient to an efficient equilibrium.[8]

Such theories, however, assume a conclusion that the literature to date does not support—that the "right" amount of external enforcement is

---

    7.   Phillip Reny may have been the first to point out the counterfactual problem in the context of common knowledge of strategic intent. *See* Philip J. Reny, *Backward Induction, Normal Form Perfection and Explicable Equilibria*, 60 ECONOMETRICA 627 (1992). The problem has also been noted with respect to common knowledge of the rationality of the players, as discussed *infra* note 37.

    8.   *See, e.g.*, Richard H. McAdams, *A Focal Point Theory of Expressive Law*, 86 VA. L. REV. 1649 (2000) (law may be designed to change social norms by expressing societal preference for one norm over another); Cass R. Sunstein, *On the Expressive Function of Law*, 144 U. PA. L. REV. 2021 (1996) (law provides a focal point on which players can coordinate, making achievement of equilibrium more probable).

usually very low. The assumption seems clearly at odds with informal observation of everyday life, in which law plays a nontrivial role not just in the creation of social order, but also in its maintenance. The question of the right amount of legal intervention can only be answered, in the abstract, by having a clearer understanding of what is and is not a source of difficulty in sustaining cooperation through (supposedly) self-enforcing norms of behavior. We seek to establish that the counterfactual problem opens up avenues for law that the literature has not yet identified. Most importantly, we contend, law does not enter simply to help players arrive at the cooperative equilibrium but is also required to sustain that equilibrium. This observation has the virtue of consistency with actual patterns of law enforcement.

Part I briefly describes the prior literature's approach to modeling normative behavior as a repeated game strategy profile and the attendant focus on tit-for-tat. Part II demonstrates tit-for-tat's shortcomings as a metaphor for real-world behavior, and Part III describes the def-for-dev strategy and defends it as a superior canonical strategy for discussions of social norms. Part IV describes the counterfactual problem and its (unsatisfactory) solutions within the repeated games literature. Part V identifies ways in which law can alleviate the counterfactual problem. It also demonstrates that legal sanctions that are insufficient to deter bad behavior can nevertheless increase the overall amount of cooperation in society. The appendices provide proofs of propositions stated in the text.

I
REPEATED GAMES, NORMS, AND THE THIRD-PARTY ENFORCEMENT PROBLEM

Scholarship on norms frequently references the theory of repeated games, and in particular the repeated prisoner's dilemma. The essential attribute of a repeated game is that players may condition their behavior on other players' past actions. This makes repeated games analogous to normative behavior because part of what gives norms their bite is that people remember those who violate norms and may alter their behavior towards such individuals. The possibility that one's opponents in a game (or neighbors in real life) will condition their future behavior on her past behavior may inspire that person to forego actions that are in her short-run interest.

Repeated games are typically used in the norms literature, however, as metaphors rather than models. The most prominent treatments draw on the well-known results in Robert Axelrod's tournament, in which players were paired to play a repeated game in round-robin fashion.[9] The tit-for-tat

---

    9.   ROBERT AXELROD, THE EVOLUTION OF COOPERATION (1984). *See* ROBERT C. ELLICKSON, ORDER WITHOUT LAW: HOW NEIGHBORS SETTLE DISPUTES 164-65 (1991); Robert D. Cooter,

strategy outperformed all other submitted strategies in the tournament. An analysis of the results in a two-player game, however, cannot be literally applicable to the problem of norm enforcement.[10] In a two-player game, cooperation exists, if at all, because of the threat of retaliation by the other player. Yet one of the characteristic features of norms enforcement that scholars seek to explain is third-party enforcement. In many social settings, a person who fails to cooperate is subjected to sanctions by the community at large, not simply by the person who was the direct victim of the defection. For many scholars, that phenomenon is an essential characteristic of norms.[11]

Much of the norms literature treats third-party enforcement as requiring an explanation from outside the repeated game framework. The concern is that third-party retaliation is not individually rational. Punishing norm violators benefits the entire community, yet the punisher bears the full cost of punishment. A potential punisher, then, will prefer to free ride on some other third party's efforts, creating a brand-new collective action problem. Several scholars refer to this "second-order collective action problem" as a particular challenge for economic theories of social norms.[12]

Several prominent theoretical contributions to the law and norms literature are inspired by this apparent puzzle. Richard McAdams attempts to create costless third-party punishment by positing that third parties punish by withholding esteem.[13] Because esteem can be conferred costlessly, he argues, punishment by withholding esteem is not costly and does not create a second-order collective action problem. Eric Posner, by contrast,

---

Decentralized Law for a Complex Economy: The Structural Approach to Adjudicating the New Law Merchant, 144 U. PA. L. REV. 1643, 1658 (1996).

10.   We are not the first to note this point. *See* Lewis A. Kornhauser, *Are There Cracks in the Foundations of Spontaneous Order?*, 67 N.Y.U. L. REV. 647, 666 (1992).

11.   *See, e.g.*, Richard A. Posner & Eric B. Rasmusen, *Creating and Enforcing Norms, with Special Reference to Sanctions*, 19 INT'L REV. L. & ECON, 369, 372 (1999); Eric A. Posner, *Law, Economics, and Inefficient Norms*, 144 U. PA. L. REV. 1697, 1699 (1996).

12.   *E.g.*, ELLICKSON, *supra* note 9, at 237 ("Incentives must be provided at an infinite regress of levels."); ROBERT C. ELLICKSON, THE EVOLUTION OF SOCIAL NORMS: A PERSPECTIVE FROM THE LEGAL ACADEMY (Yale Law Sch., Program for Studies in Law, Econ. & Pub. Policy, Working Paper No. 230, 1999) ("Although many legal scholars intuitively reject the Hobbesian view that order is never possible without a central authority, they have struggled to come up with a plausible explanation of how people ever are able to solve this second-order collective action problem."); Avery Katz, *Taking Private Ordering Seriously*, 144 U. PA. L. REV. 1745, 1749 (1996); McAdams, *supra* note 6, at 352 ("Even when the norm benefits the group, a *second-order* collective action problem remains: if others enforce the norm, the individual can gain the norm's benefits without bearing enforcement costs . . . ."); Eric A. Posner, *supra* note 11, at 1718 ("The prisoner's dilemma reproduces itself at every level of sanctioning behavior.").

To be sure, a number of scholars have also expressed other reasons to be skeptical of normative self-enforceability. Eric Posner, for example, highlights the likely informational problems faced by putatively self-enforcing norms, including the very serious problem of determining when a deviation has occurred, both factually and legally. POSNER, *infra* note 14; *see also* Jody S. Kraus, *Legal Design and the Evolution of Commercial Norms*, 26 J. LEGAL STUD. 377 (1997).

13.   McAdams, *supra* note 6.

recognizes that norm enforcement is costly, but argues that it creates a corresponding benefit of signaling the enforcer's reliability.[14] By signaling reliability, an individual gains opportunities for future cooperative endeavors, and that benefit outweighs the cost of norm compliance. Robert Cooter solves the second-order collective action problem by making the failure to comply with a norm (including failure to punish a norm violator) costly because norms are internalized. Thus, "[a] person who internalizes a norm may punish people who violate it, even at a cost to himself."[15] Each new approach is designed to solve the second-order collective action problem.

In many circumstances, however, there is no second-order collective action problem. On grounds of parsimony, if nothing else, before resorting to these complex add-ons—however ingenious and compelling on their own terms—it is worth exploring in more detail how one might model norm enforcement within the structure of an *N*-player repeated game. The second-order enforcement problem is well understood within repeated game theory, and common solution concepts are designed to deal with it.

Consider an *N*-player repeated prisoner's dilemma. Each player plays against all *N*-1 opponents in each round and is perfectly informed about the history of play. Now imagine that, in some round *t*, Player 1 defects with respect only to Player 2. Of course, the strategy that Player 2 has adopted might call for retaliatory defection against Player 1 in the next round. But what about the other *N*-2 players? Perhaps their strategies call for them to defect against Player 1 also. If so, we could consider those strategies to constitute third-party norm enforcement. The model is no less realistic a depiction of third-party enforcement than the two-player repeated game is of second-party enforcement.

Imagine that for some specification of the repeated prisoner's dilemma game, the following results hold. First, there is an equilibrium strategy that produces a high level of cooperation. Second, that strategy calls for defection in response to any prior round defection including those against other players. Third, the equilibrium is not merely a Nash equilibrium, but is also subgame perfect. A strategy is subgame perfect if no player has an incentive to deviate from the equilibrium strategy given any conceivable history of play, even those histories that are not reached when all play according to the equilibrium. If such an equilibrium exists, there is no second-order collective action problem, because incentives are indeed "provided at an infinite regress of levels."[16]

Is there such an equilibrium? Implicitly, norms scholars appear to believe that there is not. But this may be because norms scholars have

---

14.   ERIC A. POSNER, LAW AND SOCIAL NORMS 18-27 (2000).

15.   Robert Cooter, *Normative Failure Theory of Law*, 82 CORNELL L. REV. 947, 962 (1997).

16.   ELLICKSON, *supra* note 9, at 237. The model in Bendor & Mookherjee, *supra* note 2, meets these conditions.

focused principally on a single strategy: tit-for-tat. As we discuss in the next Part, tit-for-tat as typically described in the norms literature does not fit the bill. It is not a subgame perfect equilibrium of the *N*-player prisoner's dilemma for any discount factor. But tit-for-tat is only one of many possible strategies that can support cooperation, some of which are subgame perfect. In the subsequent Part we describe a particular strategy that has these desirable properties and that is a reasonable metaphor for the type of third-party enforcement we observe in real life.

II
BEYOND TIT-FOR-TAT

It is easy to understand why tit-for-tat has captured the imagination of norms scholars. The strategy is essentially the command to "do unto others as they have just done unto you," and thus comports well with actual beliefs and behaviors. Moreover, tit-for-tat gained the highest cumulative score in Axelrod's famous tournament, and Axelrod dedicates a fair portion of his book to explaining the attractive properties that account for tit-for-tat's victory.[17]

Notwithstanding these advantages, there are two reasons why tit-for-tat is not well-suited to answer many of the most significant questions posed by the norms literature. These two reasons correspond, in turn, to two important ways in which Axelrod's tournament differs from the kinds of settings that the norms literature studies.

A.   Tit-for-Tat for More Than Two

Tit-for-tat was devised for the interaction of two players. Axelrod's tournament, although a round-robin with many competitors, pitted strategies against each other two at a time. As discussed above, this makes Axelrod's results not strictly comparable to the multiparty interactions of interest to norms scholars.

A natural extension is to consider strategies for an *N*-player prisoner's dilemma in which each player interacts with all others in every round. Of course, the assumption that every member of the community deals constantly with every other, and for identical stakes, takes us rather far from real life. But it is, nevertheless, a considerable improvement on two-party interactions. It also provides a baseline for refinements in which one could, for example, limit interaction to overlapping circles of neighbors, impose restrictions on the information players have about "distant" agents, and so on.

---

17.   AXELROD, *supra* note 9. *See also* ELLICKSON, *supra* note 9, at 164-65; Cooter, *supra* note 9, at 1658.

Unfortunately, it is not obvious how to adapt tit-for-tat to a many-player setting—even one as simple as that just described. The defining quality of tit-for-tat is its command to cooperate in response to prior-period cooperation and defect in response to prior-period defection. The meaning of this command is clear in a two-player game. But it is ambiguous in a multiplayer game in which a player might cooperate as to some opponents and defect as to others. A simple solution is to instruct the player to "do unto others as they have just done unto you in particular, ignoring what they have just done unto third parties." Thus, if Player 1 defects only against Player 2, only Player 2 defects against Player 1 in the next round. To adopt this interpretation, however, is to effectively divide the many-player game into parallel two-player interactions, and the resulting analysis will have no more to say about third-party enforcement than a single two-party repeated game.

To be of any use, then, tit-for-tat must be extended to the N-player setting in a way that encompasses third-party enforcement. This is not as simple as redefining the strategy to call for defection in response to another player's defection against any player. Such an instruction remains ambiguous when an opponent defects against some and cooperates with others. The strategy must be further refined by giving a rule to deal with such situations, such as "defect if another player defects as to anyone else, otherwise cooperate," "defect if another player defects as to everyone else, otherwise cooperate," or "defect if another player defects as to more than half of her opponents, otherwise cooperate."

But even after we define a rule for such situations, tit-for-tat remains a troublesome model of third-party enforcement. This is because tit-for-tat, as commonly defined, does not distinguish between unprovoked defections and retaliatory defections—it instructs players to defect in response to any prior-period defection. Third-party enforcers who step into the fray and punish by retaliatory defection are to be treated as if they themselves had defected unprovoked. Punishing the punishers seems like an unpromising way to sustain third-party enforcement. We thus must further extend tit-for-tat to somehow distinguish between warranted and unwarranted defections.

B.   *The Trouble with Tit-for-Tat, However Defined*

Even if we extend tit-for-tat so that it calls for some measure of third-party enforcement, and even if we do so in a manner that does not punish third-party punishers, analysis of the resulting strategy will still send a misleading message on the problem of third-party enforcement. Tit-for-tat-like strategies—strategies based on staggered imitation—cannot sustain third-party (and even second-party) enforcement because they are never subgame perfect. As we will demonstrate, the incentive not to deviate in the

first place implies the incentive to forgive those who have just deviated. Yet to be plausible as a model of normative behavior, third-party norm enforcement must rely on credible threats.

If tit-for-tat is so flawed, how did it emerge victorious in Axelrod's tournament? This question points to the second sense in which Axelrod's tournament is inapplicable to settings of concern to the norms literature. In that tournament, strategies were inputted into a computer and then run without allowing programmers to intervene midcourse.[18] This enabled the programmers to precommit to threats that the threatened would have expected the programmer to reconsider had the programmer been able to change the program in the middle of the run.

1. Two-Player Tit-for-Tat

Tit-for-tat's enforcement problem arises even in a two-player setting. We will show this by establishing that what is required of discount factors to prevent unprovoked defection will imply that players want to forgive their defecting opponents. Formally, we prove that tit-for-tat can never be subgame perfect at any discount factor.

Consider first the temptation to deviate from tit-for-tat without provocation. Focus, without loss of generality, on the first round of play. Player A will resist deviating from tit-for-tat only if no deviation is profitable. One particular deviation from tit-for-tat is to defect in the first round and then return to tit-for-tat starting with round two. This produces the payoff stream $m, s, m, s, \ldots$, where m is the "meanie" and s is the "sucker" payoff.[19]

On the other hand, if Player A does not deviate from tit-for-tat, she enjoys perpetual cooperation with Player B. Therefore, to induce Player A not to deviate from tit-for-tat, it is necessary that Player A prefer the payoff stream $c, c, c, c, \ldots$ (where c is the joint-cooperation payoff) to the payoff stream $m, s, m, s, \ldots$. This will be the case if cooperative payoffs exceed the average of meanie and sucker payoffs and Player A's discount factor is large enough to overcome the fact that, given $m > c$, the payoffs in the stream $m, s, m, s, \ldots$ are front-loaded compared to those in $c, c, c, \ldots$.

The same requirement must hold for Player B if Player B is to resist deviating unprovoked. Assuming that it does hold for Player B, let us focus on whether Player B would really follow through on tit-for-tat's threat of

---

18. *See* AXELROD, *supra* note 9.

19. In round one, this deviation yields meanie payoffs. In round two, Player A returns to tit-for-tat. Because Player B cooperated with Player A in round one, Player A cooperates with Player B in round two. Player B, however, defects against Player A in round two, giving Player A sucker payoffs in that round. In the third round, Player A defects against Player B again because Player B defected in the second round. Player B, however, cooperates in round three, because Player A was cooperative in the second round. Thus, Player A obtains meanie payoffs in round three, and we are back to where we started.

retaliation should Player A defect in round one. In conducting this analysis we assume, as required for Nash equilibrium, that Player B believes that Player A plays tit-for-tat going forward.[20] If Player B retaliates by defecting in round two, then Player A will defect against him in the succeeding round, and defection will echo back and forth between the two players. Going forward from round two, then, Player B receives the payoff stream $m, s, m, s, \ldots$.

Alternatively, Player B may deviate from tit-for-tat in round two by ignoring Player A's round one deviation rather than punishing it. By hypothesis, from round two on, Player B will always be treated as he has just treated Player A. Therefore, in round three both players will cooperate again and this will continue ad infinitum. Thus, if Player B deviates from tit-for-tat in round two by forgiving Player A's round one defection, then Player B obtains the payoff stream $c, c, c, c, \ldots$ going forward.

Tit-for-tat's threat of punishment, therefore, is self-contradictory. Inducing cooperation requires both that players prefer the payoff stream $c, c, c, c, \ldots$ to the payoff stream $m, s, m, s, \ldots$ and that players are expected to follow through on punishments. But, if players do prefer the payoff stream $c, c, c, c, \ldots$ to the payoff stream $m, s, m, s, \ldots$, then they cannot be expected to follow through on punishments. Intuitively, what is required to induce the players to keep a cooperative relationship on track also induces them to put it back on track, should it derail, without punishing the player responsible for derailing it. This establishes:

PROPOSITION I: With two players, tit-for-tat is never a subgame perfect equilibrium no matter what the discount factor.

2. *Many-Player Tit-for-Tat*

The same problem plagues any adaptation of tit-for-tat to many-player games to the extent it retains the property that individuals are to be treated as they have just treated others. To show this, we will examine a particular many-player setting and a particular version of tit-for-tat.

Suppose that $N+1$ players[21] play a round-robin prisoner's dilemma within each round. That is, each player plays the two-player prisoner's dilemma once against each of her N opponents in each round. This is repeated indefinitely, and players discount the future. All players observe the full history of play, including the past interactions of other players.

In the context of this game, consider the following extension of the tit-for-tat strategy for a two-player game, which we will call "third-party-enforced tit-for-tat": Cooperate with all players in round one. Thereafter,

---

20. There is an important and nonobvious sense in which this assumption is a questionable manner to proceed. We return to this issue in Part IV.

21. The notation turns out to be more convenient if we define the variable N to be the number of opponents rather than the number of players.

"do unto others as they have just done unto you in particular, except defect against any player that has just defected against any other player without provocation." We will say that a player is provoked to defect against an opponent only if the opponent has just defected against her or has just defected against another without provocation.[22] Notice that this strategy retains the property that individuals are to be treated as they have just treated others. Yet, the strategy does entail some amount of third-party enforcement: if Player 1 defects unprovoked as to Player 2, all players defect against her in the following round. Further, this strategy avoids something of the problem mentioned in the previous section. When Player 3 punishes Player 1 after Player 1 defects unprovoked against Player 2, Player 3 is not then treated as a defector by players 2, 4, 5, . . . , N+1. This would seem to give the strategy at least a fighting chance of sustaining credible third-party enforcement. Nonetheless, as we show in an Appendix,[23] the strategy fails for essentially the same reason as does two-player tit-for-tat:

> PROPOSITION 2: Consider the N+1-player repeated round-robin prisoner's dilemma. No matter what the discount factor, third-party-enforced tit-for-tat is not a subgame perfect Nash equilibrium of this game.

III

AN IMPROVED MODEL OF THIRD-PARTY ENFORCEMENT: "DEF-FOR-DEV"

So far, we have noted that community enforcement of norms should be based on credible threats, and we have demonstrated that tit-for-tat's threat of enforcement is not credible. Fortunately, some of the infinitely many possible strategies for the repeated prisoner's dilemma are subgame perfect for sufficiently large discount factors.[24] In the abstract, any such strategy could be considered a model of norm enforcement. The law and norms literature has not merely been interested in showing the theoretical possibility of self-enforcing norms, however. It has also sought to identify game-theoretic metaphors that match reasonably well with the actual behaviors we observe in well-ordered communities. In keeping with that objective, we offer an alternative to tit-for-tat as a canonical metaphor for real-world norm enforcement.

---

22. Note that all defections in round one are unprovoked. Because defecting without provocation is well-defined for round one, under the above definition it is also well-defined for round two. By the same reasoning, it is well-defined for round three, and so on.

23. *See infra* Appendix: Third-Party-Enforced Tit-for-Tat.

24. To identify one of the many possibilities, consider the so-called "grim strategy." A version of this strategy instructs players to 1) cooperate in round one, and 2) beginning in round two, defect as to all opponents if and only if any player has defected in any prior round. The grim strategy is a subgame perfect equilibrium when discount factors are high enough. However, in real life, responding to a single defection with perpetual defeetion is the exception rather than the rule (warring clans like the Hatfields and McCoys are not held out as paragons of normative behavior).

Consider a type of interaction that is common in discussions of social norms. In a particular neighborhood, an informal understanding exists that all of the houses will be painted in light, neutral colors to preserve the historic feel of the neighborhood, of which all the residents are proud. One resident, Mr. Smith, decides to paint his house a very bright purple. Neighbors react by gossiping about Smith's poor taste and uncooperative nature. They also shun him by refusing to attend a party he throws after finishing the painting. But one neighbor, Mr. Jones, goes out of his way to praise Smith, attends the party, and extends him a reciprocal invitation. Neighbors will also react negatively to Jones. They may gossip about him to punish him for siding with the original transgressor.

This simple tale illustrates a further element of normative behavior that a strategy should capture. Individuals sanction those who defect unprovoked. They also sanction those who fail to punish others' norm violations. Tit-for-tat, as ordinarily described, punishes defection with retaliatory defection. But real-world norms often also require defection in response to unwarranted cooperation. It is precisely this problem that requires Ellickson, for example, to posit the existence of both "substantive" and "remedial" norms.[25]

What a model of normative behavior requires, then, is a strategy that punishes not simply defection, but also deviation from expected behavior. We accordingly introduce and analyze a simple example of such a strategy and show that, for sufficiently high discount factors, it is a subgame perfect Nash equilibrium and therefore consistent with credible third-party enforcement.

Consider the following strategy, which we will call "defect-for-deviate," or "def-for-dev": 1) Start in round one by cooperating with all opponents; 2) Defect against a player if and only if she deviated from the def-for-dev strategy in the immediately preceding round.[26]

To see how def-for-dev operates, suppose that we have ten players, and all ten cooperate as to all others in round one. Because all have acted consistently with the strategy to date, the strategy instructs each player to cooperate with everyone in round two. If all ten do cooperate with all

---

25. *See* ELLICKSON, *supra* note 9, at 184-219.

26. Three technical notes are in order regarding the definition of "def-for-dev." First, we adopt the convention that part 2 of these instructions does not apply to round one, for which there is no immediately prior round. Second, although these instructions are recursive (they refer back to themselves), they nevertheless generate a unique well-defined strategy for each player, which by definition specifies an action in the current round after every conceivable history of play. The instructions generate this strategy inductively in the mathematical sense: they are clear about what Player i (and every other player) is supposed to do in round one; and, given what every player was supposed to do in round $t-1$ as well as what every player actually did in round $t-1$, the instructions are clear about what Player i (and every other player) is supposed to do in round t. Third, because of the one-to-one association between the instructions and the strategy they generate, there is no inconsistency in referring to both the instructions and the generated strategy as "def-for-dev."

others in round two, then all have again acted in conformity with the strat-
egy in round two, and the strategy then instructs each player to cooperate
with everyone in round three. Assume, however, that Player 1 defects in
round three against Player 8, but no other player defects. Then Player 1 has
not followed def-for-dev in the third round: she has violated def-for-dev's
instruction to defect against another player only if that player deviated
from def-for-dev in the immediately prior round. Thus, in round four, def-
for-dev tells Players 2 through 10 to defect against Player 1 and cooperate
with all other opponents. Further, it tells Player 1 to cooperate with all
opponents. Thus, Player 1 is to "play the sucker" against all opponents in
round four as punishment for her deviation in round three. Now suppose
that all players except Player 2 follow the strategy in the fourth round and
defect against Player 1 (Player 1 also follows the strategy and cooperates
with everyone). Player 2 cooperates with Player 1 in the fourth round. Then
Player 2 and only Player 2 has deviated from def-for-dev in the fourth
round. Thus, in the fifth round def-for-dev tells Player 1 and Players 3
through 10 to defect against Player 2 and cooperate with all others. Further,
it tells Player 2 to cooperate with all opponents. Thus, Player 2 plays the
sucker in round five as punishment for failing to punish Player 1 in round
four. And so on.

A.   Def-for-Dev's Simplicity

As is evident from its succinct definition, def-for-dev is as simple to
describe as two-player tit-for-tat. It is also simple for the players to imple-
ment. Players need only carry over two pieces of information for each op-
ponent from round to round. They must remember what each opponent did
in the last round and what each was supposed to do.[27] This is no more in-
formation than is required to implement the version of many-player tit-for-
tat considered earlier.[28] There, players had to know not only whether a
player had defected in the last round, but also whether that defection was
unprovoked. All else equal, the simpler a strategy, the more plausible it is
that players will actually adopt it.

B.   Def-for-Dev's Credible Enforceability

Although def-for-dev shares tit-for-tat's simplicity, it does not share
tit-for-tat's flaw of relying on non-credible threats. Suppose that each
player believes that all other players are playing def-for-dev. To say that
def-for-dev is a Nash equilibrium is to say that it is in each player's interest

---

27.   Technically, def-for-dev is a Markovian strategy when states are defined along two
dimensions: what players did and what they were supposed to do. For a formal definition of strategic
simplicity, see Dilip Abreu & Ariel Rubinstein, *The Structure of Nash Equilibrium in Repeated Games
with Finite Automata*, 56 ECONOMETRICA 1259 (1988).

28.   *See supra* Part II.B.2.

to play def-for-dev herself. To say that def-for-dev is not just a Nash equilibrium, but also subgame perfect, is to say that it is also in each player's interest to punish those who deviate from def-for-dev. Indeed, subgame perfection further implies that it is in each player's interest to punish those who fail to punish those who deviate, as well as those who fail to punish those who fail to punish those who deviate, and so on. In other words, to show that def-for-dev is a subgame perfect equilibrium is to show that it is capable of satisfying the seemingly daunting requirement that "incentives . . . be provided at an infinite regress of levels."[29]

In an Appendix[30] we show that def-for-dev is a subgame perfect equilibrium if players do not discount their future payoffs too heavily:

> PROPOSITION 3: Consider the N+1-player repeated round-robin prisoner's dilemma. Def-for-dev is a subgame perfect Nash equilibrium for this game if all players' discount factors exceed the greater of $(m-c)/(c-s)$ and $(d-s)/(c-s)$.

Just how "patient" do the players need to be? As explained in the Appendix, depending on the relative size of m, c, d, and s, the patience required to insure that def-for-dev is a subgame perfect equilibrium may actually be less than that required to insure that tit-for-tat is a mere Nash equilibrium.

Assuming players are this patient, the intuition behind def-for-dev's subgame perfection is straightforward. If play has been cooperative so far, then def-for-dev instructs all players to continue cooperating in the current round. In deciding whether to follow this instruction, each player weighs the short-term net benefits of meanie payoffs in the current round, $m - c$, against the discounted loss of sucker payoffs in the next round, $\delta(c - s)$. The player finds this threat credible, because any opponents who fail to punish her in the next round not only forego meanie payoffs in that round, but also will themselves be forced to play the sucker in the round after next. The condition that the discount factor also exceed $(d - s)/(c - s)$ is needed to insure that deviators are willing to play the sucker for one round (thus losing $d - s$) in return for the discounted benefit of mutual cooperation in the following round, $\delta(c - s)$.

C.   Other Attractive Properties of Def-for-Dev

Aside from its simplicity and its potential for credible enforcement, def-for-dev has several other properties that make it intuitively attractive as a model of norm enforcement.[31] First, def-for-dev is retaliatory, more so than tit-for-tat. Both strategies treat the unprovoked deviator the same,

---

29.   ELLICKSON, *supra* note 9, at 237.
30.   *See infra* Appendix: Subgame Perfection of Def-for-Dev.
31.   *See generally* AXELROD, *supra* note 9 (listing intuitively attractive properties of norms including "retaliation" and "forgiveness").

defecting against her in the following round. But tit-for-tat, as noted above, does not typically distinguish between provoked and unprovoked defection.[32] It requires retaliation against each. Def-for-dev, on the other hand, refrains from punishing a player who punishes in the manner called for by the norm. It does, however, punish those who fail to punish when they should. Def-for-dev, accordingly, punishes those who are not team players.

Second, def-for-dev is more forgiving than tit-for-tat. In def-for-dev, if the deviator takes her lumps in the subsequent round, play returns to perpetual cooperation going forward. Alternatively, the player may wait to perform this act of contrition as long as she likes. In the meantime, she is "ostracized": she remains in a relationship of mutual defection until she chooses to take her penance and play the sucker. Deviant players will, in fact, find it in their interest to take their punishment and resume cooperating as soon as possible. Compare this to tit-for-tat. By virtue of its odd echoing property, tit-for-tat never gives players a chance to forgive. It is true that the punishing player forgives the defector, but just as he does so, the defector is punishing the punisher for his punishment, and so like the figures in a broken cuckoo clock, the two players continue to hit each other over the head in alternating fashion, ad infinitum.

Def-for-dev, then, is a significant improvement over tit-for-tat as a tool for analyzing the behavior of rational utility maximizers. But our analysis is still not a sufficient basis for concluding that cooperation is self-sustaining. Unfortunately, problems more intractable than that of third-party enforcement lurk within the repeated games literature.

IV
THE SELF-CONTRADICTORY NATURE OF SELF-ENFORCEMENT

To say that a strategy S is a Nash equilibrium of a game is to note that it is a best response to itself.[33] To say, however, that the players will actually play S requires a set of statements about the players' beliefs. In particular, each player must expect that all others will adopt S.[34] If we consider a pair of cars approaching one another on a narrow road to be a game, the strategy "pass on the right" is a Nash equilibrium. That means only that passing on the right is the best strategy for each player if each thinks that the other will do the same. If the players have no common understanding of what the other will do, dumb luck is their only hope of avoiding a collision.

---

32. *See supra* Part II.A.

33. A more general definition accommodates Nash equilibria in which the players do not play identical strategies. *See* FUDENBERG & TIROLE, *supra* note 3, at 11-23, 84.

34. For a more precise account of the epistemic conditions for equilibrium play, see Robert Aumann & Adam Brandenburger, *Epistemic Conditions for Nash Equilibrium*, 63 ECONOMETRICA 1161 (1995); B. Douglas Bernheim, *Rationalizable Strategic Behavior*, 52 ECONOMETRICA 1007 (1984); David G. Pearce, *Rationalizable Strategic Behavior and the Problem of Perfection*, 52 ECONOMETRICA 1029 (1984).

We ignore the question of how players come to have a common understanding of strategic intent, other than to note that models of communication, learning, and evolution in games attempt to describe plausible conditions under which players' beliefs would converge on a particular strategy.[35] Instead, we discuss a paradox involving the maintenance of such an understanding in settings like the repeated games considered in the norms literature.[36] We argue that the existence of this paradox challenges the notion that norms can ever be "self-enforcing." This, in turn, leads us to identify several ways in which law plays a vital supporting role in the enforcement of social norms.

A.   The Counterfactual Problem

Consider how tit-for-tat and def-for-dev operate in the repeated prisoner's dilemma. Both ensure cooperation by threatening to punish uncooperative behavior. Thus, assuming a sufficiently high discount factor and a collection of rational players who believe that all others will play def-for-dev, the sequence of play is uncomplicated: everyone cooperates at all times. The threat of retaliatory defection is never put to the test because no one defects. But cooperation is sustained only because each player believes that the others would retaliate in response to a deviation.

What happens when, contrary to def-for-dev's instructions, someone defects? At one level, the answer is simple. Like all strategies, def-for-dev gives a complete recipe for action after any observed history of play, including those histories that the strategy rules out. Thus, every other player knows to defect against the deviant in the next round and forgive her in the round after next, and the deviant knows to accept the punishment and then resume cooperation.

Although each player knows how to respond given her choice to play def-for-dev, that choice was based on the belief that all other players would play def-for-dev also. As we have just noted, were that belief true, the defection that now requires punishment would not have happened. The paradox is this: should the original belief, on which the decision to play this strategy is based, persist in the face of inconsistent evidence?[37]

---

35.   *See, e.g.*, Chris William Sanchirico, *A Probabilistic Model of Learning in Games*, 64 ECONOMETRICA 1375 (1996).

36.   The problem we will describe applies to any situation in which one player observes another's past play before making a choice that depends on her own beliefs about the other's future play.

37.   *See* Reny, *supra* note 7, at 627 (proposing a modification to sequential equilibrium whose motivation "stems from the difficulty in justifying sequentially rational behavior in subgames reachable only through a violation of sequential rationality"). Selten, who is credited with devising the notion of subgame perfection (for which, in part, he became a Nobel Laureate) also wrestled with the counterfactual problem for a time. *See* Reinhard Selten, *The Chain Store Paradox*, 9 THEORY & DECISION 127, 138 (1978).

Game theorists have paid greater attention to a related problem, involving common knowledge of players' rationality, as opposed to common knowledge of players' strategy choices. *See* KEN BINMORE,

This paradox is a manifestation of the "problem of the counterfactual" in epistemology.[38] Is it logically consistent for an individual to have a belief system that extends to what will happen in states of the world that her belief system rules out? The problem is particularly acute as applied to the theory of games—known to some philosophers as "interactive epistemology"—wherein an individual's belief system concerns the actions and belief systems of other agents, who in turn have similar belief systems that include the individual herself. One philosopher/game theorist describes the resulting confusion for both players and analysts:

> If a player were to find herself at an information set with which the theory of the game she is using is inconsistent, she would be deprived of a theory upon which to base her decisions. This would leave the other players (and the game theorist) without a theory, too, since they would become unable to predict what she will do, and would therefore be unable to decide what to do themselves.[39]

B.   Implications for the Self-Enforceability of Norms

The problem of the counterfactual is not just a logical conundrum. It calls into question whether norms such as def-for-dev can really be self-enforcing in any meaningful sense. When the objects of individuals' beliefs are the actions and beliefs of others, each individual can generate a counterfactual scenario for the others by deviating from their expectations. For any norm—however much "in equilibrium" or "perfect"—this is a serious source of instability.

To illustrate, imagine that two players are playing an indefinitely repeated prisoner's dilemma. Adopt the conventional hypothesis that the players' rationality[40] and the structure of the game[41] are common knowledge between these players.[42] Imagine, provisionally, that it is also

MODELING RATIONAL PLAYERS: PARTS I AND II, ECONOMIC GAMES, BARGAINING AND SOLUTIONS 85-167 (Omar F. Hamouda & J.C.R. Rowley eds., 1997); Kaushik Basu, *Strategic Irrationality in Extensive Games*, 15 MATHEMATICAL SOC. SCI. 247 (1988); Philip J. Reny, *Rationality in Extensive-Form Games*, 6 J. ECON. PERSP. 103, 112 (1992).

    38.   *See* DAVID LEWIS, COUNTERFACTUALS (1973).

    39.   Cristina Bicchieri, *Self-Refuting Theories of Strategic Interaction: A Paradox of Common Knowledge*, 30 ERKENNTNIS 69, 70 (1989).

    40.   Players are rational if they choose their strategies to maximize their expected discounted payoffs, where their expectations are determined by their probabilistic beliefs about what other players will do. *See* Bernheim, *supra* note 34; Pearce, *supra* note 34.

    41.   The structure of the game is described by the players' payoff functions. A player's payoff function is a mapping whose domain is the set of strategy choices for both players and whose range is the real line. The parameters of the payoff function in this case include the one-shot prisoner dilemma payoffs and the discount factors.

    42.   A proposition is "common knowledge" among a group of individuals if all know the proposition to be true, all know that the others know the proposition to be true, all know that all others know that all others know the proposition to be true, and so on. For example, if you are wearing blue shoes and we are facing one another in broad daylight, then the fact that you are wearing blue shoes is common knowledge between us. By contrast, if someone merely told me that you are wearing blue

common knowledge that the players have adopted the def-for-dev norm. Suppose, however, that one of the players, call her "Player D" or "D," has just defeeted unprovoked in period ten.

The other, conforming player, call him "Player C" or "C," will face the following considerations. According to dcf-for-dev, Player C should defcct against Player D in period eleven and then resume cooperating with her in period twelve. But D's defeetion should cause C to question his belief that he and D had a common understanding about the norm in operation. After all, were such an understanding in place, it would not have been in D's interest to defect in period ten.

In short, Player C's belief that Player D is playing def-for-dev has been falsified. Accordingly, C must formulate some new belief about the norm D is following in order to predict her future actions. In particular, C will wonder what understanding of the norm D could have had that would have made it in D's interest to defect in period ten.

As it happens, there is no shortage of subgame perfect equilibria that start off cooperatively but then defect in period ten.[43] One possibility is that Player D believed she had the "right" to defect with impunity in period ten. As we demonstrate in an Appendix,[44] a class of subgame perfect equilibrium strategies exists that we call "def-for-dev-with-defection-rights."[45] Such strategies are like def-for-dev, except occasionally one player or the other has the right to defect with impunity. Any attempt to retaliate for the exercise of that right is punished by several periods of mutual defection before play returns to def-for-dev.[46] In the abstraet, one might still consider def-for-dev the more plausible and attraetive norm. But of the two, only

---

shoes, I might be aware of the proposition, but you might not know that I am aware, nor would I know whether you know that I am aware.

43.   On a teehnieal note, this eonclusion does not, strictly speaking, follow from the folk theorem for repeated games. The folk theorem says only that any payoff in the eonvex hull of the set of stage game payoffs that exceeds each player's minimax payoff can be produced as the long-term average discounted payoff of a subgame perfect equilibrium of the repeated game for sufficiently patient players. *See* FUDENBERG & TIROLE, *supra* note 3, at 150-60. The folk theorem does not say that any path of play corresponds to a subgame perfect equilibrium of the repeated game. Abreu's work on repeated games is more relevant. *See* Dilip Abreu, *On the Theory of Infinitely Repeated Games with Discounting*, 56 ECONOMETRICA 383 (1988). Abreu shows essentially that any path of play is that of a subgame perfect equilibrium if and only if it can be enforced by "simple punishments," which essentially entail punishment by reversion to the deviating player's least favorite subgame perfect equilibrium going forward. *Id.* at 384-85. Abreu does not, however, characterize what these worst subgame perfect equilibria look like. In general, their strueture will depend on the size of discount factors.

44.   *See infra* Appendix: Def-for-Dev-with-Defection-Rights.

45.   The Appendix: Def-for-Dev-with-Defection-Rights, *infra*, defines these strategies in the more general context of N-player games.

46.   The punishment for deviating during def-for-dcv segments of play is still that one must play the sucker in a subsequent round. But such punishment is postponed if it would fall during a period in which the non-deviator has defection rights anyway. See the Appendix: Def-for-Dev-with-Defection-Rights, *infra*, for details.

def-for-dev-with-defection-rights is consistent with D's actions in period ten.

Imagine that, after reflection, the "period ten defection right for Player D" norm emerges as most probable in Player C's mind following D's period ten defection. Given this revision of C's beliefs, what action does C take in period eleven and thenceforth? In a sense, this question has already been answered. By hypothesis, C believes this new norm to be in place, and it requires that C cooperate in period eleven despite D's period ten defection. Because C cares enough about maintaining a relationship that is mostly cooperative, C chooses to cooperate in period eleven. Thus, even though the original plan called for C to respond with a period eleven defection in the counterfactual that D defected in period ten, the relevance of that plan has indeed been countered by the facts. Accordingly, C has abandoned it for a new theory about his relationship with D, one that both explains D's defection at period ten and instructs C to cooperate in period eleven.

This might appear to be all well and good. Although cooperative play was on shaky ground for a moment, the players' relationship has apparently been saved by Player C's decision that acquiescing in Player D's defection is in his best interest. This result, however, is a serious blow to the sustainability of an equilibrium. Note that if C can reason his way to cooperation after a period ten defection by D, he can (and presumably will) do the same in response to defections in periods eleven, twelve, thirteen, and so on. Furthermore, if D correctly anticipates C's reasoning process, D will indeed defect in period ten—and in periods eleven, twelve, thirteen, and so on. In fact, if D works through the reasoning at the outset of the game, D should start defecting right from period one. But then we have made a mockery of the notion that def-for-dev is a self-enforcing norm.[47]

---

47. The analysis need not be confined to beliefs and beliefs about beliefs, as opposed to beliefs about beliefs about beliefs, etc. Nor need it be confined to period ten; nor to D's defection rather than C's; nor to the players' original belief in def-for-dev, as opposed to the beliefs that replace it following the counterfaetual defection.

Thus, we can also nest Player D's analysis of Player C's analysis in the mind of C. Maybe all that we have just said about D is actually going through the mind of C in period eleven. In other words, maybe C believes that D defected in period ten as part of an opportunistie attempt to increase D's payoffs. How should C respond? It is important to note that we cannot conclude that C would call D's bluff and defect in period eleven as per def-for-dev's original instructions. C's decision in period eleven to acquiesce in D's period ten defection does not require that C believe that D really thought she had the right to defect in period ten. C will choose to cooperate if he believes that D believes she will get away with it and is thereby prepared to defect in period twelve if C defects against her in period eleven. Indeed, it is not even necessary for C's acquiescence that C thinks D believes she can get away with it. It is suffieient that C thinks that D thinks C believes D will get away with it. The analysis continues like this ad infinitum with ever more nested belicfs about beliefs about beliefs about beliefs . . . that will also on their own suffice to induce D's defection in period ten.

Moreover, it is not diffieult to see that all we have said applies as well to all periods and with either player in the role of unprovoked defector. Indeed, all we have said also applies to def-for-dev-with-

The likelihood of such instability only increases when Player D and Player C play their game in a community that practices third-party enforcement. After D defects against C in period ten, Player C not only has to explain to himself why D would have thought it in her interest to do so, but C also has to second-guess how his neighbors will interpret the same event. Perhaps C sees D's period ten defection as nothing more than an opportunistic effort to exploit the counterfactual problem which, if properly resisted, will induce D to quietly resume playing def-for-dev in period twelve. Thus, were C and D playing alone, this scenario would show D's period ten deviation unfolding just as prescribed in def-for-dev. But in deciding what to do in period eleven, C has to take his other neighbors' beliefs into account as well. C may believe that some neighbors will ascribe a defection right to D in period ten, as described above. In this case, if C defects against D in period eleven, some community members will defect against C in period twelve as punishment for violating D's period ten defection right. Alternatively, C might be fairly sure that none of his other neighbors will really fall for D's opportunism in period ten. Nevertheless, he might suspect that many of them will act as if they did in order to justifiably sucker C in period twelve should he defect in period eleven against D. And, just as in the two-player case, D may be pondering how the community will react to a defection as she decides whether to defect in period ten.

In some sense, then, there is indeed a third-party enforcement problem. But that problem does not arise from second-, third-, or n^{th}-order incentives, given belief in the norm. Rather, it derives from the fact that the confusion in beliefs caused by unanticipated deviations is multiplied by the number of players whose belief revision must be second-, third-, and n^{th}-guessed.

This lack of stability is not just a characteristic of def-for-dev. One can tell a similar tale about any strategy profile that would be otherwise self-enforcing. Even the always-defect strategy, which is trivially a sub-game perfect equilibrium, is subject to the counterfactual problem—except there the "problem" is good news to some extent. What happens after Player D deviates from always-defect in period ten and instead cooperates? Always-defect tells Player C to defect in period eleven. C's desire to do that is based on C's belief that the norm is always-defect and that his cooperation will never be rewarded with future cooperation from his opponent. But if D had really believed that there is no cooperative norm, why would D have cooperated in period ten? Perhaps D believes that the cooperative

---

defection-rights and any norm that would replace it. Replacement norms are as vulnerable as those they replace.

Thus, as is usually the case with paradox, the result is indeterminacy and confusion. In the end, the fact that any of these norms are subgame perfect equilibria is cold comfort, and we are left with the puzzle of how any norm could ever get started, let alone sustain itself for any number of periods.

norm begins now, following the suckering she suffers in period ten. (This belief also corresponds to a subgame perfect equilibrium of the game.) Perhaps, believing that D has reasoned thus, C cooperates in period eleven. And, perhaps thinking C will reason that D has reasoned thus, D cooperates in period ten.

This description of emergent cooperation resonates with Robert Axelrod's description of how norms of nonaggression spontaneously arose between opposing troops in World War I.[48] On a daily basis, troops in both sides' trenches were engaged in a prisoner's dilemma: aggression was prudent whatever the other side's approach. But opposing troops often faced each other for long periods of time, and a live-and-let-live norm emerged in certain settings. In describing how these cooperative norms arose out of a situation that undoubtedly began as always-defect, Axelrod relates that a "key factor was the realization that if one side would exercise a particular kind of restraint, then the other might reciprocate."[49] Thus, one side would decline to shell food and water supplies behind the enemy's trenches, even though doing so was a feasible and effective tactic, and the other would respond with equal restraint.[50]

The inherent instability of subgame perfect equilibria can thus make efficient outcomes more likely in some circumstances. Always-defect is a subgame perfect equilibrium and arguably the initial condition of any repeated interaction. Were always-defect truly stable, cooperative norms would never arise. The problem, of course, is that any cooperation norm that may arise is also not self-sustaining. In the case of trench warfare, directives from both sides' high commands ended cooperation between the troops.[51] But it is not clear that such norms would have lasted in any event, for the live-and-let-live norm that emerged was vulnerable to the same instability that was its genesis.

### C.	Game-Theoretic Workarounds

Game theorists have devised several ways of dealing with this counterfactual paradox that seem to leave intact the self-enforcing character of subgame perfect equilibria. In fact, however, these "solutions" are more ways of ignoring the problem than of solving it.

Indeed, the most common approach of game theorists has been to ignore the problem outright. Typical descriptions of equilibrium strategy profiles implicitly assume that players neither attempt to explain the deviations they observe nor attempt to predict how others would explain

---

48.	*See* AXELROD, *supra* note 9, at 73-87. Axelrod bases his analysis on TONY ASHWORTH, TRENCH WARFARE, 1914-1918: THE LIVE AND LET LIVE SYSTEM (1980).

49.	AXELROD, *supra* note 9, at 78-79.

50.	*Id.* at 79.

51.	*Id.* at 81-82.

their own deviations.[52] Thus, when Player C sees Player D defect in period ten, the conventional treatment is to assume that C ignores the fact that D would not have done so were def-for-dev really commonly understood. Similarly, it never occurs to D that defecting might cause C to reevaluate his belief in def-for-dev in a way that could benefit D.

A somewhat more introspective approach is to acknowledge the counterfactual problem but assume that it has no ramifications for future play.[53] Under this approach, after Player D defects in period ten, Player C admits to himself that his belief that both players would always play def-for-dev was mistaken. C assumes, however, that this apparent parting of the minds was a one-time affair having no bearing on future play. Furthermore, D, in contemplating whether to defect in period ten, knows that C will treat her defection in this way and thus D refrains from defecting.

Pierpaolo Battigalli argues that such an approach is internally consistent.[54] But logical consistency is not the only issue. Another issue is whether the approach fairly describes the world. And on this score we question whether players will so optimistically compartmentalize their belief in the common understanding supporting the equilibrium. Is it plausible that Player C would think def-for-dev is still viable after Player D has deviated from it 100 times in a row? Is it really plausible that C's belief in def-for-dev would remain unaffected after he is the only one in a population of 100 who played def-for-dev last period? If not, then even a single defection in a single period should lead C to assign some weight to the possibility that the common understanding surrounding def-for-dev has disintegrated.

Another common means of avoiding the counterfactual problem is to assume that players explain deviations entirely in terms of mistakes in the execution of strategies, rather than in terms of differences in players' intentions or beliefs. Thus, a deviation is ascribed entirely to an opponent's "trembling hand" as she figuratively sets down the pieces on the board.[55] Yet another means of dealing with the counterfactual problem is to assume

---

52.   *See, e.g.,* FUDENBERG & TIROLE, *supra* note 3, at 92-99.

53.   *See* Pierpaolo Battigalli, *On Rationalizability in Extensive Games,* 74 J. ECON. THEORY 40 (1997).

54.   *See id.* In his theoretical structure, players' beliefs have a lexicographic structure. That is, players have a hierarchy of beliefs. Each expects the other to play strategy *S*1 in preference to any other strategy. However, if it beeomes clear that *S*1 is·unavailable, then each expects *S*2; if *S*2 is unavailable, then *S*3, and so on. Further, we can require that the hierarchy be common knowledge. Finally, we can impose the additional modest rationality constraint that a player eliminates from consideration only strategies that are falsified by available information. Thus, when Player C observes Player D defect in round *t*, he eliminates only those strategies that require cooperation in round *t*. For example, if strategies *S*1 through *S*4 are falsified by the round *t* defection, then both players expect strategy *S*5 to prevail from round *t*+1. There are other ways of reestablishing equilibrium, with the common thread that players are assumed to update their beliefs in response to new information in predictable ways.

55.   *See* Selten, *Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games, supra* note 2 (introducing the concept of "trembling hand" perfection).

that players ascribe the deviation entirely to a mistake in their understanding of the motivations (that is, the payoff functions) of the other players. Under this approach, when Player C sees Player D defect, he might deduce that D's meanie payoffs were temporarily inflated, thus making a temporary deviation worthwhile.

Both approaches rule out by fiat the possibility that players will conclude that the deviation evidences a crack in the common understanding going forward. The logical pressure built up by Player D's period ten defection is released entirely through channels other than players' doubts about the future viability of the norm. Technically, this occurs consistently with "rationality" because it is implicitly assumed that the players put probability of exactly one on the maintenance of their common understanding going forward, while placing something less than probability one on at least one of these other norm-sustaining explanations for the defection. When the players apply Bayes' Theorem[56] to update their beliefs about the game on seeing D defect in period ten, all of the updating is allocated to those components of the game over which the players entertain some doubt. The players' beliefs in the common understanding going forward, to which they assign a prior probability one, are not updated.

To the extent, however, that one admits the possibility that players entertain some grain of doubt about the norm, then one must admit the possibility that some of the doubts that Player C has when he sees Player D deviate in period ten are allocated to C's former belief that def-for-dev is commonly understood to be the norm going forward. And once this possibility is admitted, we are back to the logical problem discussed above: genuine self-enforceability requires a common understanding that a norm is in place even after events tend to show that it is not.[57]

---

56. Bayes' Theorem is an analytical formula that provides a means of updating a probability assessment in response to new data. *See* ROBERT V. HOGG & ELLIOT A. TANIS, PROBABILITY AND STATISTICAL INFERENCE 101-03 (5th ed. 1997).

57. It is worth remarking that restricting attention to "totally mixed" strategies (strategies in which every pure strategy receives some probability weight) is also no solution. First, it would be difficult to argue that any mixed strategy profile could ever be stable in any meaningful sense. Mathematically, a rational player who mixes two pure strategies must be perfectly indifferent between the two pure strategies (and thus any probability mixture thereof). Nevertheless, a mixed strategy equilibrium requires that the rational player mix in the two pure strategies in the precise proportion that makes her opponent also completely indifferent between the pure strategies that the opponent is mixing. Requiring that all players choose one precise probability mixture among a continuum of alternatives, each of which they regard as equally attractive, seems an unlikely start for stability.

Second, while it is certainly true that there are no unreached information sets when all players use totally mixed strategies, once again internal consistency is not the only issue. If, for example, Player A observes Player B take an action which receives little weight under what she understands to be Player B's mixed strategy, why would Player A not entertain the possibility that Player B has changed strategies? In conventional game theory, the reason is as discussed in the text: players are assumed to put probability one on the event that their opponents are following the mixed strategies prescribed for them. Thus, however little weight a player puts on a particular pure strategy in her totally mixed strategy, the observing players ascribe her play of this action entirely to the unlikely event that it is the

Game theory's workarounds for the problem of the counterfactual are not justifications for ignoring the problem, they are simply ways that it can be ignored. While these approaches have had the benefit of allowing game theorists to maintain internal consistency, they are no solution for legal scholars who desire to understand what game theory really teaches about the viability of extralegal norms.

V

IMPLICATIONS FOR LAW

Starting from this updated understanding of the difficulties of applying game theory to social norms, how might law assist parties in overcoming the obstacles to cooperative behavior? First, law ameliorates the counterfactual problem by introducing an enforcer whose actions are predictable even in the face of unexpected actions by others. Second, and equally important, law helps to contain the damage done by bad actors and keep it from spilling over into interactions among "good" players. We discuss this point in some detail because it helps explain how law can maintain order without deterring all bad acts.

A. Courts and the Counterfactual Problem

The counterfactual problem is not merely a logical puzzle, but captures an important feature of real-life interactions. Norm violators often argue that they should not be punished because their violation was inadvertent or, given the circumstances, excusable. These arguments sometimes succeed. Nevertheless, we recognize intuitively that excessive leniency creates an opportunity for would-be norm violators to deviate intentionally but to claim that the deviation was unintentional or excusable. Thus, a child caught red-handed immediately (and instinctively) protests "it was an accident" or "she hit me first." Desiring not to create bad incentives, parents do not automatically credit these claims.

These common interactions illustrate the counterfactual problem—a defector may try to exploit the cooperator's desire to believe that the norm is still intact and that last period's defection was extraordinary. The cooperator is torn between hope that the cooperative norm is still in place and concern not to appear a pushover. In some circumstances, the decision whether to punish the norm violator requires that the other players agree on what motivated the deviation. In other situations, the community may need a means of determining what behavior will be tolerated going forward—does the fact that the child's sister hit first justify a retaliatory hit, or will the retaliation also result in punishment?

---

outcome of the mixed strategy that they think she is playing, never questioning whether she might have switched mixed strategies to one that gives the observed action more weight.

Law provides ongoing assistance of both sorts. Courts announce expected behavior going forward. They fill gaps in social norms by declaring that, for example, Player D indeed had the right to deviate in round ten, and therefore retaliatory defection is improper. They also adjudicate claims that a violation was inadvertent and therefore deserving of less, or no, punishment.

There is an inherent problem in relying only on decentralized community action to serve these functions in the face of the counterfactual problem. It is impossible to determine objectively whether a defection was unintentional, and accordingly there is no way to assure that all members of the community will agree on whether the defection should be punished. Similarly, if there has been no prior occasion for the community to consider whether particular circumstances excuse a deviation, we cannot be sure that the community will agree ex post on the appropriate resolution.

The best that can be done given these uncertainties is to establish a second-order norm that the determination of some person or institution (i.e., the court) will be honored. A community may be willing and able to punish informally without any intervention by a central authority. The threat of punishment, however, is not entirely credible if the community cannot easily reach consensus on when punishment is appropriate. By contrast, a community's willingness to follow a court's decision does not require that all members agree with the decision, but only that all expect all others to follow it.

Of course, this moves the counterfactual problem up a level. When we establish a second-order norm that the court's decision will be honored, how can we know what to do if another player disregards the court's decree? Accepting the premise that the norm of honoring the court's decree is incentive-compatible for each individual given that such a norm is commonly accepted in the community, then an individual's failure to honor that decree would call into question whether the norm is really commonly accepted—and then we are back to the same counterfactual perplexity that we encountered in a world without the court.

Yet, in an important respect, the situation here is not the same. The difference lies in the government's ability and willingness to force compliance irrespective of community beliefs. The possibility that the defendant in a lawsuit ignores the court's decree does not lead to the counterfactual problem because the incentive compatibility for the defendant of following the decree does not rely solely on a common understanding in the community that such court decrees are followed. Rather, the incentive compatibility of following the decree is grounded in the court's ability and willingness to punish the defendant irrespective of (or at least with less dependence upon) how the defendant's behavior might affect any common understanding among community members. The court's capacity and

resolve to back its commands with force provide a solid platform on which society can build more complex informal enforcement systems.

Thus, although law—meaning centralized commands backed by force—may constitute only a small portion of the total rule-making and enforcement system, it is a necessary one. Without it, the remaining parts of the system may unravel in the face of uncertainty about whether all players continue to expect one another (and expect that others expect, and so on) to continue to comply with the extant norms.

B.   Containing the Damage from Congenital Deviators

Even though a court is not subject to counterfactual uncertainty in the face of a defection to the same extent as community norm enforcers, the fact of the defection can still be bad news for the sustainability of coopera-tion. Imagine, for example, that the court finds that a particular defection was wholly unwarranted—the defector understood the game and its pay-offs and chose to defect nonetheless. Because cooperation was the player's best ex ante choice, given the assumed payoffs, other players' assumed strategies, and the defector's assumed discount factor, other players may resolve the counterfactual problem by concluding that the player is "bad"—that is, the player has a discount factor much lower than the other players expected. If so, other players now anticipate that this bad player will defect again.

Can a legal system like ours—one that neither detects all deviations nor punishes with sufficient severity to deter all deviations—provide any help to players confronted with a bad player? Or will the presence of unde-terred bad players destroy the cooperative equilibrium? Fortunately, even a legal system that cannot entirely deter deviation can limit the damage done by deviation, as we will show with a slightly enriched version of our basic *N*-player prisoner's dilemma model.

Imagine a group of *N* players about to commence an indefinitely re-peated prisoner's dilemma game in which each expects the others to play def-for-dev. Assume that not all players have the same discount factor.[58] Imagine as well that it is common knowledge among the current players that there is occasional turnover in the population. Members of the popula-tion are periodically replaced with other new players.[59] This is a convenient way to model the fact that the relevant community for any given individual is ever changing. Given this structure, players' discount factors account not just for their impatience, but for the possibility that they will not be dealing

---

58.   On a technical note, the analysis of equilibria up to now has not required homogeneity of players' discount factors. All that has been required is that the lowest discount factor in the population was at least the minimal discount factor needed to make all one-stage deviations unprofitable.

59.   For a similar structure, see generally Drew Fudenberg & David K. Levine, *Steady State Learning and Nash Equilibrium*, 61 ECONOMETRICA 547 (1993).

with this community in future periods.[60] Some current and incoming play-ers have such a low discount factor that it is not in their interest to play def-for-dev against the other members of the community—maybe these bad players know they are likely leaving the community very shortly, or per-haps they are simply impatient by nature. Players do not know ahead of time what the discount factor of any new entrant is.

Strictly speaking, def-for-dev is no longer an equilibrium of the game structure that we have defined. Among the set of players—broadly defined to include even those waiting in the wings—there are those who will not find it in their interest to play def-for-dev even if they know that the rest of the population will. These low discount players will prefer to play always-defect against a population of players playing def-for-dev.

But among those whose discount factors are high enough, def-for-dev may still be a best response to the full range of their potential opponents' strategies.[61] These players will choose to cooperate the first time they meet any player, and they will respond to deviation in the same manner as def-for-dev.

In this situation, however, the benefits to cooperation for these "good" players are reduced in comparison to a situation in which all N opponents always play def-for-dev. The intuition is as follows: a representative player compares the gains expected from perpetual cooperation with the gains she could achieve by defecting. The potential entry of players in the following period who will defect whether or not the representative player cooperates now reduces the gains from cooperation and so increases the discount factor required to sustain cooperation.[62] If we imagine that dis-count factors among the good players are distributed across some range, then the existence of a bad player will reduce the expected amount of co-operation among the good players. Players who might find it just barely profitable to conform to def-for-dev in a population full of conforming players may decide that always defecting is best when a sufficient portion of the population is likely to be bad.

This prompts an important observation: legal liability can increase the expected amount of cooperation among the remaining players even if it is insufficient to deter the bad player. Assume that, on average, the bad player is required to reimburse one-fourth of her meanie payoffs to her

---

60. Formally, if δ' is the discount factor reflecting the players' patience, and p is the probability that a player will be replaced next period, then the relevant discount factor is $\delta = p\delta'$.

61. We add to def-for-dev the proviso that new entrants are by convention not considered among those that have deviated in the immediately prior round and that new entrants start by cooperating.

62. The same can be said of the possibility that any entrant in the current period is impatient. To see this, consider only the payoffs derived from interaction with an impatient player. Defecting rather than cooperating against such a player in the current round increases current payoffs from sucker payoffs to defection payoffs. Moreover, defecting rather than cooperating with this player in the current round has no impact on future payoffs with respect to this player, since even under def-for-dev one essentially plays always-defect with this player beginning with the following period.

victims. This may reflect the fact that defectors are not always caught, that damages are systematically undercompensatory, or both. The bad player's subjective discount factor may be such that three-quarters of a meanie pay-off in round one is worth more than perpetual cooperative payoffs, in which case she still prefers to defect. However, the liability adds weight to the "cooperative" side of the ledger for the other players and may tip the balance in favor of a cooperative strategy. Any amount of liability for defection, then, increases the probability of cooperation among other players, even if the liability is insufficient to deter defection.

The same is equally true of incapacitation of bad players through imprisonment.[63] Incapacitation reduces the proportion of uncooperative players in the population by removing those who repeatedly defect from the population and making room for new entrants who may well be cooperative. Assuming that the number of opponents is roughly constant (perhaps there is some technological or physical constraint on the number of players with whom one can play at any single time), incapacitation of prior-period defectors increases one's average payoff in future periods. The effect of increasing the probability of cooperation among each member of the group ·increases the stability of the norm within the entire group. Liability and incapacitation can therefore reduce the likelihood of cascading defections leading to social disintegration.[64]

In summary, the legal system's interventions seem too light and episodic only when we view them solely from the perspective of a would-be defector. A focus on the decision problems of those predisposed to cooperation, by contrast, shows that a legal system with a relatively light touch can maintain social order even when the society contains some persistently bad actors. The legal system's ability to impose liability or to incapacitate defectors increases the range of discount factors over which cooperators will remain cooperative.

CONCLUSION

The use of repeated game theory in norms scholarship has, to date, been heavily influenced by the optimistic results of Axelrod's prisoner's dilemma tournament. Unfortunately, tit-for-tat, while simple, ingenious,

---

63. One could imagine a more complex model in which players play two types of prisoner's dilemma games—a low-stakes game in which the legal system punishes defections through liability (that is, a wealth transfer to the victim), and a high-stakes game in which defections are punished by imprisonment.

64. The analysis is similar to that of an evolutionary model. We can assume some probability of defection in a given round and ask whether other players, operating on the assumption that the defection will be repeated, will find it in their interests to defect as well. Such games have a characteristic dynamic in which some proportion of defectors within the population will tip the value-maximizing action from cooperation to defection, in which case social breakdown occurs rapidly. Paul G. Mahoney & Chris W. Sanchirico, *Competing Norms and Social Evolution: Is the Fittest Norm Efficient?*, 149 U. PA. L. REV. 2027, 2047 (2001).

and successful in Axelrod's special setting, has drawbacks in an N-player context in which players carry out only credible threats. We suggest a modified strategy in the spirit of tit-for-tat that serves as an improved metaphor for real-world, multiplayer interactions.

It is not, however, the case that repeated game theory provides a compelling account of the sustainability of cooperation so long as the right strategies are considered. In a more rigorous and abstract modeling exercise, after all, one could simply enlist the entire set of subgame perfect equilibria of an N-player game. But any such equilibrium still would be subject to the counterfactual problem. Seemingly irrational behavior is a serious challenge to any solution based on an assumption of universally rational calculation.

This, we believe, is one reason that law will always lurk nearby in any discussion of social norms. And that is as it should be. Theories of social norms have added a great deal to our understanding of how cooperation can arise, as it were, outside the shadow of the law. But any plausible theory of social norms must account for the continued presence of law, even if law is not the primary instrument of social control.

APPENDIX: THE REPEATED PRISONER'S DILEMMA

A.   The Stage Game and Repeated Game

Before turning to the repeated prisoner's dilemma, we briefly discuss the prisoner's dilemma itself. The phrase "prisoner's dilemma" stands for a class of two-player, two-strategy simultaneous move games. In two-player, two-strategy simultaneous move games generally, each of two players simultaneously and in isolation chooses from one of two strategies. The combination of the players' strategy choices determines each player's payoff. Each player chooses a strategy that maximizes her expected payoffs given her probabilistic beliefs about which strategy her opponent will choose. Such a strategy is called a "best response" to the given beliefs about the strategy of one's opponent. A player's probabilistic beliefs may indicate that she is sure that another player will play a particular strategy. In this case, we say that a best response to such beliefs is a best response to the particular strategy itself.

In a prisoner's dilemma, in particular, one of the strategy choices for each player is usually labeled "cooperate" and the other "defect." Moreover, the payoffs for each combination of the players' strategy choices have two properties: 1) whatever each player's beliefs about her opponent's strategy choice, her best response is to defect (this is expressed by saying that defect is a "dominant strategy"); and 2) both players' payoffs are higher when both choose to cooperate than when both choose to defect.

Figure 1 shows payoffs that yield these two properties. Player A's strategy choice determines the row, Player B's the column. In each cell of the table, Player A's payoffs are listed first and Player B's second.

FIGURE 1

	Cooperate	Defect
Cooperate	10, 10	0,15
Defect	15,0	5,5

More generally, if we use variables as in Figure 2, prisoner's dilemma payoffs satisfy the three inequalities: $m > c > d > s$. (As a mnemonic for the strategy permutations that produce them, Axelrod calls m and s "meanie" and "sucker" payoffs, respectively.) Although not formally part of the definition, it is usually assumed that $2c > m + s$. Were this not the case, it would be socially optimal for one party to cooperate and the other to defect, rather than for both to cooperate.

FIGURE 2

	Cooperate	Defect
Cooperate	c, c	s, m
Defect	m, s	d, d

Properties 1) and 2) of the prisoner's dilemma payoffs produce the chief interpretive lesson of the prisoner's dilemma. The pursuit of individual interest (which in this case leads each player to defect) does not necessarily produce the best social outcome (because both players would be better off if they both chose to cooperate).

A game may be "repeated," meaning that the players play the identical game (referred to as the "stage game") multiple times in sequence. In the repeated prisoner's dilemma, players meet once each "round," for some prespecified number of rounds, to play the prisoner's dilemma as just described. The full history of play to date is common knowledge between the players going into each round. Thus, each player may condition her current choice in any given round on how her opponent has played to date. Equivalently, we may imagine that players choose at the outset whether to cooperate or defect in each round for each possible history of play up to that round. This full time- and history-contingent choice is also called a "strategy." So as not to confuse this with use of the same term in the stage game, the choices "defect" and "eooperate" are, in the repeated context, referred to as "actions." A list of strategies, one for each player, is called a "strategy profile."

A strategy is a hyper-comprehensive, contingent plan of action. It specifies the action a player will take whenever it is her turn to move, whatever she has observed of the course of play to date. It even specifies actions for histories that the player is certain will not obtain given her beliefs about what others will do. Of course, her choice of an action to take in response to an action she is sure her opponent will not take does not directly affect her payoffs. Because such a state of the world receives zero probability weight, it also contributes zero to the player's expected payoffs. Rather, the importance of specifying what the player would do should her opponent surprise her comes from the fact that in equilibrium theory (discussed below)—wherein players are assumed to know their opponent's strategy—her strategy plays a dual role. Not only is it her plan of action, but it also represents her opponent's beliefs about how she will and would behave. And how she would behave were they to do something other than what she expects is directly relevant to *their* choice problem. We return to this point in our discussion of subgame perfection below.

In selecting a strategy, players are assumed to account for both current and future payoffs. Usually, however, future payoffs are time discounted so that current payoffs weigh more heavily in the player's choice. Time discounting is accomplished by application of a discount factor, a number between 0 and 1.[65] For any given round t, the payoffs one period hence in $t+1$ are multiplied by this factor (usually represented by δ) to make them comparable to round t payoffs. Payoffs two periods hence are multiplied by δ^2 in comparing them to round t payoffs, and so on. The closer the discount factor to one, the more "patient" the players, or the more they weigh the effect of their current action choices on future as opposed to current payoffs.

Most commonly, the prisoner's dilemma is repeated for an infinite number of rounds. This is not to be taken literally. Rather, the discount factor stands in both for time preference and for the probability that the current round is the last. The latter interpretation of the discount factor is intuitive: if I believe there is a 50% chance the game will end after the current round, then a potential payoff in the next round is worth, at most, half as much to me as the identical payoff in the current round. Under this interpretation, an infinitely repeated prisoner's dilemma with a discount factor strictly less than one may be viewed as an "indefinitely" repeated game.

To the extent that my opponent is following a strategy that conditions her future action choice on my own current action choice, the consequences of my current action choice extend beyond the immediate prisoner's dilemma payoffs described in Figures 1 and 2, and include future

---

65. One can also represent the same kind of time discounting with a "discount rate," which operates like an interest rate in time discounting. For any given discount rate r, the equivalent discount factor is $\frac{1}{1+r}$.

payoff consequences as well. Most saliently, if defecting rather than cooperating in the current round would induce my opponent to defect rather than cooperate in future rounds, then current defection may no longer be my payoff maximizing choice taking into account both current and future payoffs. Thus, a strategy calling for current defection, given the history of play, would not necessarily maximize payoffs. Thus, property 1) of the prisoner's dilemma does not carry over to any individual play of that game in the repeated context.

B.   Nash Equilibrium

Given any game, whether a one-shot prisoner's dilemma, a repeated prisoner's dilemma, or any other, a Nash equilibrium is a strategy profile with the property that each player's strategy in that profile is a best response to the other players' strategies in the profile.

In the one-shot prisoner's dilemma there is only one Nash equilibrium: both players defect. This follows from the fact, discussed above, that defecting is a dominant strategy. Given that the other player is defecting, defecting maximizes one's payoffs ($d > s$). In contrast, cooperating is not a best response to itself: given that one's opponent cooperates one can do better by defecting ($m > c$). In other games, there may be no dominant strategy, but there may be one or more Nash equilibria.[66] As discussed in the text, if one imagines driving along a narrow road as a game, neither the strategy "pass on the left" nor the strategy "pass on the right" is dominant. However, either is a Nash equilibrium because if Player A chooses the strategy, it is in Player B's best interests to do the same and vice versa.

In a repeated prisoner's dilemma, so long as the discount factor is not too close to zero, there are infinitely many Nash equilibria. Always-defect is one such equilibrium; in general the unconditional repetition of any equilibrium in the stage game (here: both defect) is also an equilibrium for the repeated game. Tit-for-tat is another equilibrium. The full time- and history-contingent strategy that is tit-for-tat is easy to describe: 1) In round I cooperate; 2) In every other round, choose the action that your opponent chose last round. Grim-like strategies and def-for-dev, as defined in the text and notes, are also Nash equilibria.

C.   Subgame Perfect Nash Equilibrium

Any given strategy profile produces a "path of play." This is the observed sequence of actions that results when both players play their strategy in the strategy profile. For example, if both players adopt tit-for-tat, the path of play is that both players cooperate in all rounds. Because tit-for-tat

---

66.   If we allow for "mixed strategies," in which players choose randomly among the possible actions with specified probabilities, there will always be at least one equilibrium.

is an equilibrium, the path of play for tit-for-tat is also referred to as its "equilibrium path."

There are many possible histories of play that never obtain for any given strategy profile. For example, if both players play tit-for-tat, it would never be the case that both players had defected in all odd numbered rounds up to round 13. When the strategy profile is an equilibrium, we say that such histories are "off the equilibrium path."

If a strategy profile is a Nash equilibrium, then each action choice it specifies along the play path must be a best response to the opponent's strategy going forward. However, it is not required for Nash equilibrium that, should the players find themselves off the equilibrium path, the action choices prescribed by the strategy profile are a best response to the opponent's continuation strategy. Put more simply, players are free to believe that their opponents will make choices against interest off the equilibrium path.

An informal example illustrates the point. Suppose two countries, A and B, border one another and have a history of border skirmishes. Country B declares that it will react to any incursion by A's soldiers—however slight—by launching an all-out nuclear strike on A. In response, country A makes an identical declaration. Each country has a sufficient nuclear arsenal to guarantee annihilation of the other even if attacked first. Going forward, it is in the best interest of each country never to cross the other's border, given belief that the other would respond with nuclear war. Thus, the equilibrium path of play would never include an incursion. Imagine, however, that an incursion just occurred—a single Country A soldier crossed the border. Will B really retaliate by launching a nuclear war that will assure B's destruction as well as A's? The conditions for Nash equilibrium do not require that we answer this question. Equilibrium analysis assumes that the players will follow their presumed strategies.

This feature calls Nash equilibrium into question as a viable solution concept. The issue is whether it is realistic for a player to believe that her opponent will act against interest should she deviate. This issue is often referred to under the rubric of "credible threats." In the norms literature, this problem arises in an *n*-player setting and is referred to as the third-party enforcement problem.

The problem, however, is easy to solve. One simply requires that a strategy be a best response going forward for all histories of play, even those not reached by the strategy profile. This requirement is embodied in the definition of a "subgame perfect Nash equilibrium," an equilibrium that also satisfies this additional property.

D.   Extension to N-*Player Settings*

All that has been said is readily generalized to games in which there are more than two players. In this case, each player's payoffs are a function of her own choice and all the choices of her opponents. To be a best response, a strategy must be a best response to the joint choice of one's opponents.

Although some games are inherently defined for more than two players, given any two-player game there are many ways to extend that game to *n* players. The repeated prisoner's dilemma, for example, may be extended to many players by specifying that each player plays all others in pairwise fashion within each round, her payoffs in each round being the sum of the payoffs from each pairwise encounter. In this case, one may simplify by assuming that each player makes one action choice per round to be employed in each encounter in that round. This is the simple structure we adopt in the text. Alternatively, one may assume that one meets each player in each round, but chooses differently for each pairwise encounter. This is an atypical assumption, but it may also be analyzed using the same concepts. Yet another alternative is to assume that each player is randomly matched with one or more opponents in each round.

APPENDIX: THIRD-PARTY-ENFORCED TIT-FOR-TAT

PROPOSITION 2: Consider the *N*+1-player repeated round-robin prisoner's dilemma.[67] No matter what the discount factor, third-party-enforced tit-for-tat is not a subgame perfect Nash equilibrium of this game.

Consider first Player 1's temptation to defect unprovoked. One particular deviation is to defect against all opponents in the first round and then return to third-party-enforced tit-for-tat starting in round two. This deviation yields meanie payoffs against all opponents in the first round. In the second round, Player 1 returns to third-party-enforced tit-for-tat. Because no player in the prior round either defected against her or defected against another unprovoked, she cooperates with all. However, because she defected without provocation in the first round, all her opponents defect against her in the second round, giving her *N*-fold sucker payoffs. In the third round, she defects against all opponents because they defected against her in round two. Her opponents, however, cooperate with her, because she was cooperative in round two. She obtains meanie payoffs against all opponents in round three, and we are back to where we started. Thus, a one-time all-opponent deviation from third-party-enforced tit-for-tat produces the same oscillation between meanie and sucker payoffs that one observes

---

in a two-player game, except that the oscillation is now N-fold. The payoff stream from the one-time all-player deviation is now Nm, Ns, Nm, Ns, ...

In order to be sure that Player 1 will resist this deviation, the payoff from so deviating cannot exceed what Player 1 can expect to obtain if she follows third-party-enforced tit-for-tat. Following third-party-enforced tit-for-tat produces perpetual cooperation with all N players. Therefore, Player 1—and likewise all players—must prefer the payoff stream Nc, Nc, Nc, Nc, . . . to the payoff stream Nm, Ns, Nm, Ns, ...

Now suppose that Player 1 deviates only against Player 2 in round one. Consider Player 3's decision whether to defect or cooperate with Player 1 in the second round. In a manifestation of third-party enforcement, third-party-enforced tit-for-tat instructs Player 3 to defect against Player 1 in the second round: Player 1 has just deviated against another player without provocation. Player 3 expects that Player 1 will play third-party-enforced tit-for-tat from round two on. Thus, if Player 3 defects against Player 1, as instructed, she and Player 1 will end up oscillating ad infinitum between meanie and sucker payoffs. All other players, 2, 4, 5, . . . , $N+1$, will ignore Player 3's round two defection against Player 1 because Player 3's defection against Player 1 was provoked. Similarly, all of Player 3's future defections against Player 1, as the pair oscillates between defection and cooperation, will also be ignored by all players outside the pair, because each defection by Player 3 will follow a defection by Player 1 and so will be provoked. Therefore, Player 3 remains in a perpetually cooperative relationship with all opponents other than Player 1 from round two on. Player 3's payoff stream from round two on is thus $m + (N-1)c$, $s + (N-1)c$, $m + (N-1)c$, $s + (N-1)c$, ...

Alternatively, Player 3 might deviate from third-party-enforced tit-for-tat in round two by forgiving Player 1 for her first round defection. If so, Player 3 receives cooperative payoffs in round two with Player 1, and in all subsequent rounds she and Player 1 continue to cooperate. The other players do not punish Player 3 for failing to defect against Player 1 in round two or in any other round. Player 3's payoff stream from round two on is, therefore, $c + (N-1)c$, $c + (N-1)c$, $c + (N-1)c$, $c + (N-1)c$, ...

Therefore, Player 3 will only be willing to follow third-party-enforced tit-for-tat's instruction and defect against Player 1 if she prefers $m + (N-1)c$, $s + (N-1)c$, $m + (N-1)c$, $s + (N-1)c$, . . . to $c + (N-1)c$, $c + (N-1)c$, $c + (N-1)c$, $c + (N-1)c$, ... Because her relationship with Players 2, 4, 5, . . . $N+1$ is the same in both cases, this requires that Player 3 prefer m, s, m, s, . . . to c, c, c, c, ...

We have already established, however, that Player 3 must prefer N streams of c, c, c, c, . . . to N streams of m, s, m, s, ... She must, then, prefer one stream of c, c, c, c, . . . to one stream of m, s, m, s, ... Therefore, our many-person adaptation of tit-for-tat has essentially the same internal

contradiction as two-player tit-for-tat. Tit-for-tat's property of staggered imitation ensures that the incentive not to deviate unprovoked implies the incentive to forgive deviators.

APPENDIX: SUBGAME PERFECTION OF DEF-FOR-DEV

This Appendix proves that def-for-dev has no third-party enforcement problem, as that problem is conventionally understood. Specifically, we prove:

> PROPOSITION 3: Consider the N+1-player repeated round-robin prisoner's dilemma.[68] Def-for-dev is a subgame perfect Nash equilibrium for this game if all players' discount factors exceed the greater of $(m-c)/(c-s)$ and $(d-s)/(c-s)$.

Because discount factors must always be less than one, satisfying the discount factor condition in this proposition requires that $(m-c)/(c-s) < 1$, or, equivalently, that cooperative payoffs exceed the average of meanie and sucker payoffs.

Compare the discount factor condition in this proposition to that necessary to insure the weaker result that tit-for-tat is a mere Nash equilibrium. (We have seen that tit-for-tat is never subgame perfect.) That condition is that players' discount factors exceed the greater of $(m - c)/(c - s)$ and $(m - c)/(m - d)$.[69] The first portions of both discount factor conditions are the same: discount factors must exceed $(m - c)/(c - s)$ in either case. The second portions of the conditions are not unambiguously ordered. However, if cooperative payoffs plus defection payoffs are less than the sum of meanie and sucker payoffs ($c + d < m + s$), then def-for-dev's second condition is easier to meet than tit-for-tat's.[70] In this case, def-for-dev requires less patience to be a subgame perfect equilibrium than tit-for-tat requires to be a mere Nash equilibrium.

A.   The One-Stage Deviation Principle

Before proving that def-for-dev is a subgame perfect equilibrium for sufficiently high discount factors, we note a helpful principle that greatly simplifies the task. In order to test whether a strategy is subgame perfect, it

---

68.   *See supra* Part II.B.2.

69.   *See* AXELROD, *supra* note 9, at 207.

70.   The hypothesis is that $\frac{m-c}{m-d} > \frac{d-s}{c-s}$. Cross multiplying, we can rewrite this inequality as $(m - c)(c - s) > (m - d)(d - s)$. Expanding yields $mc - c^2 - ms + cs > md - d^2 - ms + ds$. Eliminating $- ms$ from both sides and regrouping terms, we have $m(c - d) + s(c - d) > c^2 - d^2$. If we expand the right hand side, we get $m(c - d) + s(c - d) > (c + d)(c - d)$. Dividing through by $c - d$ produces the condition stated in the text: $m + s > c + d$.

This condition is not logically inconsistent with satisfaction of any of the three conditions on δ mentioned in the text. The only question of inconsistency arises with respect to the first condition for both tit-for-tat and def-for-dev: $\delta > \frac{m-c}{c-s}$. Given that $\delta < 1$, this requires that $\frac{m-c}{c-s} < 1$, or $m + s < 2c$. Because $c > d$, we may simultaneously satisfy both $m + s < 2c$ and $m + s > c + d$.

suffices to test whether it does as well as all one-stage deviations from the strategy after every conceivable history of play (including those that it does not reach).[71] A one-stage deviation after a particular history of play is an alternative strategy that differs from the candidate strategy only in the current round. Thus, it is not necessary to test the strategy against all conceivable deviations, almost all of which will involve taking different actions in many different rounds under many different contingencies.[72]

The intuition behind the one-stage deviation principle is simple. Certainly, immunity from one-stage deviations after all histories of play is a necessary condition for subgame perfection: a one-stage deviation is a deviation, after all, and the candidate strategy must do better than all deviations after all histories of play. What is interesting about the one-stage deviation principle is that checking one-stage deviations is sufficient.

To see why, imagine that the prisoner's dilemma is repeated just twice and that we have determined that a particular candidate strategy is immune from one-stage deviations. Could it ever be the case that a two-stage deviation from this strategy was profitable? That would mean that deviating from the candidate strategy only in round one was not profitable, but that, if this deviation were combined with a subsequent deviation, the combined deviations would give larger payoffs. This would in turn imply that deviating later added value to the round one deviation. But that in turn would imply that deviating in round two is profitable on its own in some contingency, contradicting the hypothesis that no one-stage deviation from our candidate is profitable.

Were the game instead three rounds long, then after analogously determining that a deviation beginning in round two must be profitable on its own, we would have to conclude that this deviation was two rounds long, thus extending into period three: by hypothesis, a deviation in only round two would not be profitable. But this would mean that the round three portion of the deviation adds value, which again would contradict the assumption that no one-stage deviation is profitable. One can accordingly use induction to extend the one-stage deviation principle to games with any finite number of rounds. Indeed, the principle also applies to games with an indefinite number of rounds when players discount future payoffs. This is

---

71.   *See, e.g.*, Abreu, *supra* note 43, at 390-91.

72.   On a technical note, one might imagine that there is a corresponding one-stage deviation principle for Nash equilibrium that considers one-stage deviations only at reachable information sets. This, however, is not the case, as may be shown by counterexample. Thus, it is actually easier to determine whether a strategy is a subgame perfect Nash equilibrium than to determine whether it is a non-subgame perfect Nash equilibrium. If a strategy fails the one-stage deviation test just discussed, it may still be an equilibrium. And determining whether it still is will generally require testing it against all multiround deviations. To be sure, it is relatively easy to show that tit-for-tat is a Nash equilibrium, even though it is not a subgame perfect equilibrium. *See, e.g.*, AXELROD, *supra* note 9, at 207-08. But this is only because of tit-for-tat's special Markovian structure. In general, determining whether a strategy is a Nash equilibrium, when it is known not to be subgame perfect, is a difficult task.

because every infinitely long deviation can be approximated (in its effect on the discounted stream of payoffs) by a deviation of finite length, to which the finite-game logic just explicated will apply.

B.   *Subgame Perfection*

Let us then apply the one-stage deviation approach to def-for-dev in order to establish that it is subgame perfect with sufficiently patient players. This means that we have to rule out one-stage deviations at all conceivable histories of play. Fortunately, all possible histories may be considered en masse in two logically exhaustive categories.

1. Histories of play in which Player i did not deviate at $t-1$ (inclusive of the null history): $\delta > (m - c)/(c - s)$ is necessary and sufficient to rule out one-stage deviations.

Suppose that Player i conformed to def-for-dev at time $t-1$. Let $C \leq N$ be the number of i's opponents who also conformed to def-for-dev at $t-1$, implying the existence of $N - C \geq 0$ deviators at $t-1$.

If i conforms to def-for-dev again at t, she cooperates against each of the C period $t-1$ conformers and defects against each of the $N - C$ period $t-1$ deviators. All opponents cooperate with i. Thus, i obtains $Cc + (N - C)m$ at t. Thereafter, all players cooperate, because all conform to def-for-dev at t. The payoff stream is $Cc + (N - C)m, Nc, Nc, \ldots$.

If i deviates from def-for-dev at t, she either defects against some number of the C period $t-1$ conformers or cooperates against some number of the $N - C$ period $t-1$ deviators, or some combination of both. Because a deviating player is treated the same under def-for-dev no matter how many players she has deviated against in the last round, the most attractive deviation from def-for-dev in the current stage is to defect against all players (including last period deviators and last period cooperators). Because all players cooperate with i at t, this yields payoffs of Nm. Since i deviates from def-for-dev at t, all defect against her in $t+1$. She, on the other hand, cooperates with all at $t+1$. This yields Ns. In all subsequent rounds, there is mutual cooperation. The payoff stream is $Nm, Ns, Nc, Nc, \ldots$.

Thus, given C, the deviation is unprofitable if and only if $Cc + (N - C)m + \delta Nc > Nm + \delta Ns$. The left hand side of this inequality is decreasing in C. The right hand side is independent of C. Therefore, the deviation is unprofitable for all $0 \leq C \leq N$, if and only if $Nc + \delta Nc > Nm + \delta Ns$, which reduces to $\delta > \frac{m-c}{c-s}$.

2. Player i is one of last period's deviators: $\delta > (d - s)/(c - s)$ is necessary and sufficient to rule out one-stage deviations.

Suppose that Player i deviated from def-for-dev at time $t-1$. Let $C \leq N$ be the number of i's opponents who conformed to def-for-dev at $t-1$, leaving $N - C \geq 0$ fellow $t-1$ deviators.

. If *i* conforms to def-for-dev at *t*, she cooperates against each of the *C* period *t*–1 conformers and defects against each of the *N* – *C* period *t*–1 deviators. All opponents defect against *i*. Thus, *i* obtains $Cs + (N - C)d$ at *t*. Thereafter, all cooperate, since all conform to def-for-dev at *t*. The payoff stream is $Cs + (N - C)d, Nc, Nc, \ldots$.

In the most profitable deviation from def-for-dev at *t*, Player *i* defects against all opponents, including last period deviators and last period conformers. Given that all players defect against *i* at *t*, this yields $Cd + (N - C)d$. Since *i* has deviated from def-for-dev again at *t*, all defect against her again in *t*+1. Since we are considering only a one-stage deviation, *i* cooperates with all at *t*+1. Thus, *i*'s payoffs in *t*+1 are *Ns*. In all subsequent rounds, there is mutual cooperation. The payoff stream is $Cd + (N - C)d, Ns, Nc, Nc, \ldots$.

Thus, given *C*, one-stage deviations are unprofitable if and only if $Cs + (N - C)d + \delta Nc > Cd + (N - C)d + \delta Ns$, or $Cs + \delta Nc > Cd + \delta Ns$. Rearranging yields: $\delta > (C/N)(d - s)/(c - s)$. This condition is true for all $C \le N$ if and only if it is true for $C = N$. Thus, one-stage deviations are unprofitable for all $0 \le C \le N$ if and only if $\delta > (d - s)/(c - s)$. Note that since $c > d > s$, $1 > (d - s)/(c - s) > 0$. Hence, this condition can be fulfilled by a large enough δ.

APPENDIX: DEF-FOR-DEV-WITH-DEFECTION-RIGHTS

In this Appendix, we define and analyze a class of strategies for the *N*+1-player round-robin prisoner's dilemma. These strategies are similar in flavor to def-for-dev. They differ from def-for-dev, however, in that they endow certain players with the right to defect against certain other players at certain times. Possessing such a defection right enables a player to defect in the given instance without suffering the future sanction prescribed in def-for-dev (having to "play the sucker"). The strategies within this class differ from each other in the constellation of the defection rights that they take as given. They also differ in how they punish individuals for not respecting the rights of others by, for example, not playing the sucker against the rights holder in a particular interaction.

One specific example of this class of strategies is that which grants Player D the right to defect against Player C in period ten, as discussed in the text.[73]

A.   *Preliminary Discussion*

A first attempt at incorporating defection rights into def-for-dev might simply be to prespecify defection rights for each player, and then stipulate that a player is to defect against an opponent if and only if either that

---

73.   *See supra* Part IV.B.

opponent deviated in the immediately preceding round or the player has the right to defect against the opponent in the current round. However, this approach runs into trouble in periods immediately prior to those in which a player's opponents have the right to defect against her. If some number of opponents will be defecting against the player in the following period no matter what she does in the current period, the punishment for currently deviating is reduced.

One means of dealing with this issue is to have the opponent postpone punishing the deviator until the next time that the opponent will not be defecting against her in any event. This is the approach we take here. Periods in which an opponent has the right to defect against the player are effectively skipped over in the operation of def-for-dev with respect to that opponent. Put another way, the punishment for a deviation is "carried forward" like an unusable loss in the tax law.

How then are defection rights themselves enforced? How do we prevent a player from refusing to play the sucker when her opponent has the right to defect against her? Once again, the defection rights system is kept separate from the def-for-dev system. Defecting against a rights-holding opponent is punished by subjecting the player to several periods of mutual defection against all opponents. During the punishment period for rights violations, punishments for "normal play" deviations are carried forward.

B. *Formal Definition of Def-for-Dev-with-Defection-Rights*

In any given period, any given player may have the right to defect against some number of her opponents. The full specification of these rights is a parameter of the norm. For simplicity, we will assume that defection rights are always one-sided: no two players ever have a simultaneous right to defect against each other.[74]

Given an exogenous pattern of defection rights, we define def-for-def-with-defection-rights with an algorithm. The algorithm instructs each player i how to play against each opponent j in each period t. Player i's instruction for how to play against j in period t depends upon whether either player is being punished for a prior rights violation, whether either player has a defection right against the other, and whether player i is carrying forward a defection against j from a previous period. Whether the first and third of these "states" obtain in the current period is in turn determined by application of the algorithm to previous periods. Portions of the algorithm that set these states for future periods are typeset in italics. To initialize the algorithm, we stipulate that in the first period no players are being

---

74.   We also structure the strategy so that a deviator does not lose her right unilaterally to defect in future periods. Therefore, we also carry forward Player j's punishment of Player i when i has a defection right against j. The strategy can be modified to punish deviations with the loss of future defection rights.

punished for a prior rights violation and no player is carrying forward a defection against any other.

1. If either you or *j* are being punished for a prior rights violation, then
   a. defect against *j*.
   b. *if you carried forward a defection against* j *into the current period, you carry it forward into the next period as well.*[75]
2. If neither you nor *j* are being punished for a prior rights violation, then
   a. if you have a defection right against *j* in the current period, defect against *j*.
   b. if *j* has a defection right against you in the current period, then
      i. cooperate with *j*.
      ii. *if you fail to cooperate with* j *in this case, you will be punished for a prior rights violation for the next* r *periods, where* r *is a parameter of the norm.*
   c. *in either case 2.a. or 2.b., if you carried forward a defection against* j *into the current period, you carry it forward into the next period as well.*
3. If neither you nor *j* are being punished for a prior rights violation, and neither you nor *j* have a defection right against the other in the current period, then
   a. if you carried forward a defection against *j* into the current period, defect against *j*.
   b. if you did not carry forward a defection against *j* into the current period, cooperate with *j*.
   c. *in either case 3.a. or 3.b., and whatever choice you make in either case, you do not carry forward a defection against* j *into the next period.*
   d. *if you violate either 3.a. or 3.b., all opponents carry forward a defection against you into the next period.*

C. Subgame Perfection

We show here that def-for-dev-with-defection rights is a subgame perfect equilibrium for sufficiently large discount factors and sufficiently long punishment phases for rights violators so long as, as usual, cooperative payoffs exceed the average of meanie and sucker payoffs, and, in addition, defection rights are sufficiently infrequent and intermittent.

PROPOSITION 4: Consider the $N+1$-player repeated round-robin prisoner's dilemma.[76] Suppose that cooperative payoffs exceed the average of meanie and sucker payoffs ($c > \frac{m+s}{2}$). Suppose also that given any two Players i and j, the fraction of periods in which Player i has the right to defect against Player j is less than $\frac{c-d}{c-s}$ and the number of consecutive periods in which either has a defection

---

75. Carrying forward a defection is not a choice for Player i, but rather a means of accounting for histories of play.

76. *See supra* Part II.B.2

right against the other is never greater than some fixed bound.[77] Then, if the period r for the punishment for rights violators is long enough and the players' discount factors arc high enough, def-for-dev-with-defection-rights is a subgame perfect Nash equilibrium.

By the one-stage deviation principle,[78] it suffices to check that no one-stage deviation is profitable after any conccivable history of play. We parse the set of conceivable histories and conceivable deviations into three sub-sets.

1. Periods in which Player i is not being punished for a prior rights violation; deviations that do not entail defecting against an opponent who has a defection right against i.

Suppose that Player i conforms to def-for-dev-with-defection-rights in the current period and thereafter. Some number of her opponents $0 \le V \le N$ are being punished for rights violations in the current period. She defects against them; they defect against her. Of the remaining opponents, i has the right to defect against some number $0 \le R_0 \le N - V$, and some number $0 \le R_1 \le N - V - R_0$ have the right to defect against her. Player i suckers the R_0 players in the first group and is suckered by the R_1 players in the second group. Set $R = R_0 + R_1$.

Of the remaining players, Player i carried forward a defection into the current period against some number $0 \le D \le N - V - R$. Of these, some number $0 \le D_0 \le D$ did not carry forward into the current period a defection against Player i, and some number $0 \le D_2 \le D - D_0$ did. Player i suckers the D_0 players in the first group and mutually defects with the D_2 players in the second group. Of the remaining $C = N - V - R - D$ players, some number $0 \le C_0 \le C$ did not carry forward a defection against Player i into the eurrent period, and some number $0 \le C_1 \le C$ did. Player i mutually cooperates with the C_0 players in the first group and is suckered by the C_1 players in the second group.

Thus, a conforming player's current payoffs are

$$(D_0 + R_0)m + C_0c + (V + D_2)d + (R_1 + C_1)s.$$

Now consider the set of one-stage deviations for Player i in which she defects against no player having the right to defect against her in the current period. Because Player i is treated the same in future periods no matter what her current deviation within this class, her most attractive deviation is

---

77. More formally, the two (logically independent) conditions in this sentence arc as follows.

First, $q_{ij} \equiv \lim_{t \to \infty} \frac{1}{t} \sum_{\tau=1}^{t} x_\tau (i, j) < \frac{c-d}{c-s}$, where $x_t (i,j) = 1$ if i has the right to defect against j at t,

and $x_t (i,j) = 0$ otherwise. Second, there exists $b_{ij} < \infty$, such that if either i or j have a defection right against the other in each of the consecutive periods t, $t+1$, $t+2,\ldots$, $t+n$, then $n \le b_{ij}$.

78. *See supra* Appendix: Subgame Perfection of Def-for-Dev, pt. A.

to defect against all players lacking the right to defect against her in the current period, yielding current period payoffs of

$$(D_0 + R_0)m + C_0 m + (V + D_2)d + (R_1 s + C_1 d).$$

The current period payoff advantage from this deviation is, therefore,

$$C_1(d - s) + C_0(m - c) \geq 0.$$

This is the current gain from defecting against non-rights-holders with whom i would have cooperated.

In order for this one-stage deviation to be unprofitable, it must be that the negative future consequences of deviating outweigh this current gain. How do Player i's future payoffs change if she so deviates? Every opponent j will defect against i in the next period in which j is not being punished for a rights violation and neither i nor j has a defection right against the other. Some opponents may already be carrying forward a punishment for a prior deviation by i, and will defect anyway. But at least $C = C_0 + C_1$ players would have cooperated with i on the next such encounter. Instead, they will defect against i. Given that this is a one-stage deviation, i will cooperate with them. Thus, as a lower bound on total future losses from deviating, Player i loses $c - s$ once against each of C opponents at some point in the future. Given a discount factor strictly less than one, the negative future consequence of this lower-bound loss depends on the structure of defection rights and, in particular, the extent of the resulting delay in punishment for the current deviation. Nevertheless, the delay is uniformly bounded from above by assumption. Specifically, for each j of these C players, either i or j will have a defection right against the other for no more than b_{ij} periods in a row going forward from the next period. Therefore, as the discount factor approaches one, the negative future consequence of the deviation converges uniformly across the full subset of histories and deviations to $C(c - s)$.

Thus, there will exist a sufficiently high δ to make this kind of deviation unprofitable in terms of overall payoffs so long as the following strict inequality holds:

$$C_1(d - s) + C_0(m - c) < C(c - s).$$

Depending on the relative size of $d - s$ and $m - c$, the left hand side is at a maximum over this subset of histories when either (a) $C_0 = N$ and $C_1 = 0$ or (b) $C_0 = 0$ and $C_1 = N$. In case (b), the inequality will always hold. Case (a) requires the familiar condition

$$N(m - c) < N(c - s) \Leftrightarrow m - c < c - s.$$

2. Periods in which Player i is not being punished for a prior rights violation; deviations that do entail defecting against an opponent who has a defection right against i.

Now suppose that Player i's deviation does entail the violation of someone's defection right (and possibly other forms of deviation as well). Thus we are considering only histories in which there is some right to be violated: i.e., $R_1 > 0$. The most profitable deviation of this kind is that in which Player i defects against all players, since she will be treated the same, no matter how extensive her deviation, so long as it includes a single rights violation. In this case, her current payoffs are

$$(D_0 + R_0)m + C_0 m + (V + D_2)d + (R_1 + C_1)d.$$

Thus, the current payoff increase from deviating is

$$(R_1 + C_1)(d - s) + C_0(m - c) \leq N \max(d - s, m - c).$$

The future consequence of deviating is having to suffer mutual defection for the following r periods against all players. How bad this is depends on what one's payoffs during these r periods would have been if one had conformed in the current period. This, in turn, depends on the flow of defection rights in the ensuing r periods. Let X be the average payoff (across opponents and periods) over these r periods if the player does not deviate. Then, given r, the present value of the future loss approaches $rNX - rNd$ uniformly over this subset of histories as δ approaches I. Thus, we can find a discount factor close enough to I such that no deviation of this sort is profitable at this kind of history so long as

$$rNX - rNd > N \max(d - s, m - c)$$

which reduces to

$$r(X - d) > \max(d - s, m - c).$$

The question, then, is whether this holds for some r large enough. The answer is yes, so long as the limit of X as r goes to infinity is strictly larger than d. Let q_{jk} be the long term average fraction of periods in which j has a defection right against k. Let $q = \max_{j,k} q_{jk}$. As r goes to infinity, X will be no smaller than $\frac{N - V}{N}(qs - (1 - q)c) + \frac{V}{N}d$. (Note that for players in V, the proportion of the next r periods in which they are not still being punished goes to zero as r goes to infinity.) The relevant condition then is

$$\frac{N - V}{N}(qs - (1 - q)c) + \frac{V}{N}d > d,$$

or equivalently,

$$q < (c - d)/(c - s)$$

which was assumed.[79]

---

79.   Although this completes our treatment of periods in which Player i is not being punished for a prior rights violation, there is technically yet a third form of deviation at these periods. Player i conceivably might deviate by not exercising her own defection rights. But since this reduces her current payoffs with no effect on how she is treated in the future, such a deviation is never profitable.

3. Periods in which Player *i* is being punished for a prior rights violation.

Player *i* has no interest in deviating (cooperating) since this lowers current payoffs and has no effect on future payoffs.