The Evolution of Social Norms in Common Property Resource Use

By RAJIV SETHI AND E. SOMANATHAN*

The problem of extracting commonly owned renewable resources is examined within an evolutionary-game-theoretic framework. It is shown that cooperative behavior guided by norms of restraint and punishment may be stable in a well-defined sense against invasion by narrowly self-interested behavior. The resource-stock dynamics are integrated with the evolutionary-game dynamics. Effects of changes in prices, technology, and social cohesion on extraction behavior and the long-run stock are analyzed. When threshold values of the parameters are crossed, social norms can break down leading generally to the lowering of the long-run stock, and possibly to its extinction. (*JEL* C72, D62, Q20)

The absence of private property rights in economically valuable resources is commonly held by economists to be a primary cause of overexploitation leading to degradation and, in some instances, to complete resource extinction. Garrett Hardin's "tragedy of the commons" metaphor has been a particularly influential vehicle for the promulgation of this view. A central premise underlying this thesis is the assumption that human behavior is driven by a particular, narrowly defined conception of self-interest: the degree of resource exploitation undertaken by each individual is assumed to be that at which marginal private material gains are brought into equality with the marginal costs of extractive effort. Since the commons are characterized by a negative externality whenever the resource is scarce, this gives rise to inefficiently high levels of extraction, possibly high enough to exceed the maximum sustainable yield, and threaten thereby the long-run viability of the resource. The resulting policy implications may take two forms: either the privatization of the commons, or their appropriation or regulation by the state. The former is presumed to lead to the internalization of the externality; the latter to the enforcement of restraint by "mutual coercion, mutually agreed upon" (Garrett Hardin, 1968 p. 1247).

A number of recent studies have challenged the validity of the tragedy-of-the-commons metaphor as a general characterization of social behavior when applied to local commons such as forests, pastures, and inshore fisheries. The evidence comes from several thousand case studies by scholars in a variety of disciplines, only a few of which can be mentioned here.[1] In a study of sea tenure in the Bahia region of northern Brazil, John Cordell and Margaret A. McKean (1992 p. 191) identify a complex system of ethical codes "far more binding on individual conscience than any government regulations could ever be," which serve to ensure both sustainable aggregate har-

* Sethi: Department of Economics, Barnard College, Columbia University, 3009 Broadway, New York, NY 10027; Somanathan: Department of Economics, Emory University, Atlanta, GA 30322. For financial support Somanathan thanks the Research Training Group in Positive Political Economy at Harvard/MIT and Sethi the College of Arts and Sciences of the University of Vermont. For helpful comments and discussion, we thank Alberto Alesina, Abhijit Banerjee, Duncan Foley, Bill Gibson, Herb Gintis, Suzi Kerr, Eric Maskin, Abu Rizvi, Larry Samuelson, Robert Stavins, Steven Tadelis, Martin Weitzman, and four anonymous referees. Sethi also acknowledges valuable discussions and correspondence with Graciela Chichilnisky on some of the issues considered here.

[1] See Fenton Martin (1989, 1992) for a comprehensive two volume bibliography.

vests and an equitable distribution of access to the resource. Failure to follow the established codes of conduct can result in a variety of sanctions ranging from cultural isolation to the sabotage of equipment. Somanathan (1991) describes a variety of institutional arrangements designed to enable Himalayan villagers in India to exploit their common forests and pastures sustainably. These vary widely from village to village and range from informal customs that restrain resource use, backed only by social disapproval of violations, to government-recognized councils that appoint watchmen. The Maine lobster fisheries studied by James M. Acheson (1993) are another example of informal management of commons. There are territories claimed by groups of fishermen which are not open to outsiders. Violators are warned and their equipment destroyed if the violation continues. Such retribution is illegal and usually carried out secretly by an individual fisherman. The average annual catch has shown no trend in the last quarter of a century and the catch-to-effort ratio has been stable over the same period, which suggests that current practice is sustainable. Numerous other examples of successful and unsuccessful management of common pool resources are documented, for example, in Elinor Ostrom (1990) and in the volume edited by Daniel W. Bromley (1992).

The general pattern emerging from such studies is that restraint in the use of the commons is enforced by communities through means ranging from a total reliance on norms, to more centralized enforcement mechanisms which involve some kind of local self-government. Even in the latter cases, however, social norms have an important influence on behavior. For example, rules have force even when the formal penalties for breaking them are often very low compared to the benefit an individual obtains from noncompliance (Ostrom, 1990 Chapter 3). Clearly, one reason for this is that failure to obey the rules constitutes the violation of a social norm.

In addition to the case studies, experimental work shows that such cooperative behavior extends to controlled laboratory environments. Ostrom et al. (1992) find that even in clearly specified finite-horizon games designed to mimic the commons setting, high levels of cooperation can be sustained if the possibility of pre-game communication (covenants) is present, with or without the possibility of costly sanctions (swords). Noncooperative behavior on the part of others is sanctioned frequently by experimental subjects even when such sanctions are costly to impose, strongly suggesting that those who sanction are motivated by factors other than an exclusive concern with their own material interest.

This paper attempts to bridge the gap between economic theory and observations from the field. It provides a theory of why norms of behavior that restrain the use of common pool resources can persist even when evolutionary pressure selects against behaviors that yield lower material payoffs. It provides an explanation for the well documented observation that cooperation today will in general be followed by cooperation tomorrow, while defection today will generally be succeeded by defection tomorrow. The theory takes explicit account of disequilibrium dynamics, and sets out conditions necessary for the stability of an equilibrium composition of behavioral rules in the population. The effects of changes in prices and other key parameters on the stability of behavioral rules are explored, along with the implications for the long-run resource stock of such shifts. It is shown that even temporary shocks to the parameters can lead to permanent declines in the resource stock. The analysis makes clear why external changes can result in communities destroying resources they had previously carefully husbanded. Such an understanding of the logic of cooperative resource use can serve as a guide for deciding when external intervention is really called for, and of the form it should take. The theory of restraint we develop is consistent both with the importance of custom as documented in the case studies, and with the central premise of economics since Adam Smith that individuals will not indefinitely allow profit opportunities to pass them by. The method used is quite general and offers insights into the problem of social cooperation in a variety of situations characterized by negative externalities.

Partha Dasgupta (1993 pp. 208–11) has identified three approaches that previous work has taken to reconcile the apparent incongruity between the norm-guided restraint and sanctioning behavior that is often observed in local commons, and the assumption of self-interested behavior underlying economic theory. One approach involves the recognition of rural communities as miniature states with their own "established structure of power and authority" (p. 208) and the capacity to coerce individuals into accepting specific behavioral rules and paying for their enforcement. The scope of this explanation is limited by the fact that it cannot account for the spontaneous and decentralized sanctioning of violators that characterizes numerous local commons, particularly when such sanctioning is explicitly prohibited by the formal legal framework in the society at large.

A second approach retains the postulate of self-interest and appeals to the possibility of equilibrium strategies which can sustain cooperation under threat of sanctions in a dynamic setting. For example, Graciela Chichilnisky (1994 p. 854) attributes the persistence of cooperative resource use in local commons such as fisheries and forests to the repeated nature of the interaction among members of a small and stable group. With common knowledge of rationality and payoffs, the theory of infinitely repeated games provides for cooperation under such circumstances whenever discount rates are sufficiently low or the probability of future interaction is sufficiently high. Optimizing models of the commons as a dynamic game with an endogenous resource stock confirm that efficient resource use is a possible outcome when players are allowed to use history-dependent strategies (see Jess Benhabib and Roy Radner [1992] and the references cited therein).[2] One problem with this approach is the considerable theoretical indeterminacy which characterizes models of repeated games played by self-interested agents. The Folk Theorem (see, for example, Drew Fudenberg and Eric Maskin, 1986) has established that models of this type typically have a large number of equilibria. *Any* sequence of behavior by the players, subject only to the requirement that the average payoff obtained by each player be at least as large as that which could be obtained by defecting perpetually, is a possible outcome provided players value the future sufficiently. For example, there are many equilibria in which periods of cooperation alternate with periods of defection. Yet what is observed is that when communities achieve cooperation in resource use such cooperation is generally long lasting. Conversely, in cases when cooperation breaks down it is very difficult to recover.

The third approach (Dasgupta, 1993 p. 209) relies on the internalization of social norms through "communal living, role modeling, education, and through experiencing rewards and punishments." Once internalized, norms can provide an independent motivation for behavior that can rival and sometimes supersede self-interest. This explanation is incomplete in that it fails to establish why one set of norms is internalized at the expense of another. This paper can be seen as providing a framework for establishing which norms can be internalized.

We first focus on the case of totally decentralized enforcement and show that even here, a cooperative norm of behavior can be stable. In the concluding section we indicate how the analysis extends to more centralized systems. In the model individuals may cooperate by restraining their levels of resource use. If they fail to do so they may be sanctioned by other individuals. Punishment of this type is costly to inflict (as well as to suffer). The act of sanctioning is entirely voluntary and there are no sanctions imposed on those who fail to sanction others. In a finite-horizon setting, punishments will never be inflicted by rational, self-interested agents because once a violation has occurred, a punishment imposes a cost but confers no gain to the punisher. Instead of supposing, however, that individuals are guided by such self-interested calculation, we adopt the evolutionary-game-theoretic device of assuming that the proportion of individuals choosing a particular behavior increases when the payoff to that behavior exceeds the average payoff in the population, and decreases when the reverse is true. Hence behavior that does badly from the point of view of the individual gets weeded out, while behavior that does well is imitated. In this context, a behavior is *stable*

[2] Prajit K. Dutta and Rangarajan K. Sundaram (1993) show that overexploitation is not inevitable even if strategies are Markovian (dependent only on the current value of the resource stock).

if its adoption by most individuals is sufficient to ensure that it will not be destroyed by such evolutionary pressure.

We show that whenever there is a stable noncooperative equilibrium (one in which individuals do not restrain their use) with a positive resource stock, then there exists a *cooperative* equilibrium with a *higher* stock level. If the damages from sanctions are sufficiently high relative to the gains from defection, then this equilibrium is stable. The gains from defection depend positively on the price of the resource and the efficiency of the harvesting technology and negatively on the opportunity cost of labor. Consequently, a fall in the damages from sanctions, a rise in the price of the resource relative to that of labor, or an improvement in technology can all cause long-standing patterns of cooperative resource use to dissolve. This will lead to a reversion to unrestrained exploitation and a fall in the long-run level of the stock (perhaps even to the point of extinction) with an associated diminution of welfare. The incursion of outsiders immune to local sanctions can have the same effect. A historical instance of this is discussed in Section III.

The evolutionary approach offers insights into problems of cooperation in a wider variety of settings than that discussed here. For example, the extent of civic or cooperative behavior in a society is often discussed by political scientists in the context of an infinitely repeated game in which cooperation and noncooperation are possible equilibria. It is commonly observed (see, for example, Robert D. Putnam 1993) that such behavior is very persistent in societies. Where "civicness" is present it tends to persist, and where it is not, it does not easily come into being. But, as mentioned earlier, repeated game theory allows for any sequence of cooperation and defection to be a possible equilibrium. It does not explain why only equilibria which feature cooperation forever or defection forever are favored. The evolutionary framework, on the other hand, predicts precisely such persistence. When a given behavior is stable, individuals who adopt other strategies receive lower payoffs, and evolution (via cultural transmission and learning) weeds out the mutant behaviors.

The paper is organized as follows: Section I sets up the Common Pool Resource game as a static problem in which changes in the stock of the resource are ignored. This makes clear the nature of the problem of cooperation. Section II discusses evolutionary dynamics in this context and shows that noncooperative behavior can always be a stable outcome while cooperative behavior can be stable under certain conditions. The effects of changes in relative prices, in technology, in social cohesion, and in excludability of outsiders on cooperative behavior is analyzed. Section III takes explicit account of the effect of resource use on the stock of the resource. By integrating the dynamics of the resource stock into the evolutionary dynamics, the stability of cooperative and noncooperative behavior is considered together with the sustainability of resource use. Section IV indicates how the analysis extends to more centralized systems and concludes.

I. The CPR Game

We begin by adapting a static model of the commons as an n-person game (Dasgupta and Geoffrey Heal, 1979; Chichilnisky, 1994), which has been used in experimental work by Ostrom et al. (1992). We follow the latter in referring to it as the Common Pool Resource (CPR) game. A fixed number n of individuals have complete rights of access and removal to a natural resource (for example, fish) from a "common pool."[3] The labor or effort expended per unit of time by agent i on resource extraction is denoted by x_i and the aggregate labor expended, X, is the sum of individual labor flows:

$$X = \sum x_i .$$

[3] We use the term "common pool" to describe any resource to which multiple users have joint access, and reserve the term "common property" to refer (in Sections III and IV) to a particular property-rights regime involving explicit and/or implicit rules for the management of such resources (Siegfried V. Ciriacy-Wantrup and Richard C. Bishop, 1975).

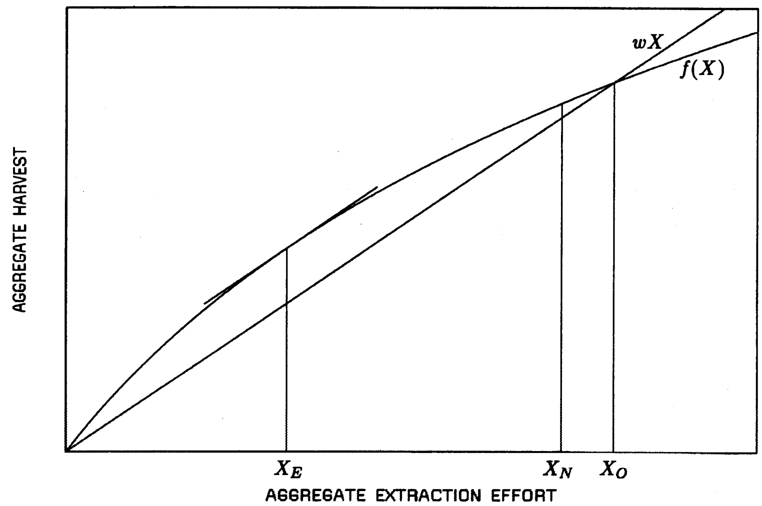

FIGURE 1. EFFICIENT, NASH EQUILIBRIUM, AND OPEN-ACCESS LEVELS OF RESOURCE EXPLOITATION

The total *stock* of the resource in existence is denoted by K. The aggregate harvest per unit of time is a function $H(X, K)$ of the aggregate labor flow X and the existing resource stock K. The total fish catch per unit of time H will be an increasing function of effort as well of the size of the stock. For a given stock, the additional catch from an extra unit of labor will clearly decrease as the total labor expended increases. Also, the higher the level of the stock, the higher will be the marginal increase in the aggregate harvest from an extra unit of labor. This is captured by the assumption that for all $X \geq 0$ and $K \geq 0$, the following hold (the subscripts denote partial derivatives):

$$(1) \quad H_X > 0, \quad H_K > 0, \quad H_{XX} < 0,$$

$$H_{XK} > 0, \quad H(0, K) = H(X, 0) = 0.$$

In this section and the next, we abstract from the effects of changes in the stock and assume that $K = K_0$, an exogenously given constant. Write $f(X) = H(X, K_0)$. The aggregate catch f now looks like a standard increasing and concave production function with labor as the input.

The cost of labor, w, which may be a hiring cost or an opportunity cost, is constant and exogenously given. The value of the aggregate harvest is assumed to exceed the total effort cost, wX up to some level X_O of total effort, and to fall below it thereafter. Normalizing the price of the resource to unity, this means that the harvest function $f(X)$ cuts the wX line from above at X_O as in Figure 1.

The share of the total harvest obtained by agent i is directly proportional to the share of agent i's effort in total effort, so that agent i's net benefit from resource extraction, denoted by π_i, is

$$\pi_i(x_1, \dots, x_n) = \frac{x_i}{X} f(X) - wx_i.$$

Therefore, the aggregate payoff $P(x_1, \dots, x_n)$ satisfies

$$P(x_1, \dots, x_n) = \sum_{i=1}^{n} \pi_i = f(X) - wX.$$

Let X_E be the level of aggregate effort which maximizes P. This is the efficient level of effort (we are ignoring changes in the stock) at which the marginal product of labor equals the wage

$$(2) \qquad f'(X_E) = w.$$

Due to the concavity of f, X_E is unique (see Figure 1). Let x_E denote the corresponding average level of individual effort: $nx_E = X_E$.

Now let us examine what the outcome will be under the usual assumption of rational, self-interested behavior by each agent. To do this, first rewrite the payoffs in terms of the *average product* of labor $A(X) \equiv f(X)/X$. In fisheries, this is called the catch-to-effort ratio. The concavity of f means that $A(X)$ is decreasing. As more labor is expended in catching fish, the catch per unit of labor falls. The average product exceeds the wage for small X and falls below it for $X > X_O$. These facts are clearly seen in Figure 1, where for any X, $A(X)$ is the slope of the line joining the origin to $f(X)$. Agent i's payoff is now

$$\pi_i(x_i, X) = x_i(A(X) - w)$$

while the aggregate payoff is

$$P = X(A(X) - w).$$

If the resource is characterized by *open access*, so that the number of users can expand without limit, then it is clear that labor will be put in until the average product equals the wage and rents are driven to zero (H. Scott Gordon, 1954; Martin L. Weitzman, 1974). The aggregate labor expended in this case is X_O (the subscript denotes open access). This is clearly inefficient.

Things are not quite as bad in the situation analyzed here (with a fixed n), but there will still be overexploitation. To see why, note that at the efficient level X_E, labor is expended until the point when the extra rent $A(X_E) - w$ from an additional unit of labor is balanced by the loss $X_E A'(X_E)$ resulting from the consequent fall in the average product of labor. However, from the individual's point of view, while the extra rent from an additional unit of labor is still $A(X_E) - w$, the loss is only $X_E A'(X_E)/n$, which is *less* than the loss to the group as a whole. Consequently, each individual will put in more labor than is efficient. There will still be positive rents from the resource, however, as long as n is fixed. For if rents were driven to zero, with $A(X) = w$, then it would pay any individual to reduce her labor input, thus raising the average product above the wage, and making her rents positive.

This is the classic problem of the commons: each individual would be better off if all would restrain their use, but it is never in the interest of any individual to do so. The relevant facts are illustrated in Figure 1 and may be summarized as follows.[4]

PROPOSITION 1: *The CPR game has a unique Nash equilibrium. The equilibrium is symmetric, with $x_i = x_N$ for all players i. It is inefficient and involves overexploitation with $x_E < x_N$. There are positive rents in equilibrium: $X_N \equiv nx_N < X_O$ so that $A(X_N) > w$.*

The prediction of this static model is unambiguous: there will be overexploitation of common pool resources. Numerous studies show that in fact, this is not always true. One reason is clearly that, as seen in the case studies, individuals have the option of imposing sanctions on other agents in response to their observed extraction levels. These range from social disapproval to physical damage such as the destruction of equipment. Moreover, such punishments are costly not only for the punished, but also for the punisher. It may be risky to damage someone else's equipment. Ostracizing someone involves losing the possibility of beneficial interaction, and even expressing disapproval can have similar effects. The cost of punishing someone may be low, but it will only rarely be absent altogether.

Consider a two-stage game, the first stage being the CPR game, and the second being one in which each individual can punish any of the others. When one person punishes another, we shall suppose that the former incurs a cost γ and inflicts a loss δ on the latter. In this CPR game with sanctioning, the payoff to agent i becomes

$$\pi_i = x_i(A(X) - w) - \delta k_i - \gamma l_i$$

[4] Existence of a symmetric equilibrium with overexploitation is proved by Dasgupta and Heal (1979). The proof of uniqueness is a straightforward extension which we omit. Alison Watts (forthcoming) proves uniqueness in a more general context. Proofs of all other formal results are collected in the Appendix.

where k_i is the number of agents that punish i and l_i is the number of agents that are punished by i.

Can individuals now be deterred from overexploitation by the threat of punishment? Not as long as they are all payoff maximizers. For once effort levels have been chosen, any individual who sanctions others will lower her payoff, regardless of the sanctioning behavior of the rest. Therefore at the second, sanctioning stage of the game, a payoff maximizer will never go through with a punishment. Anticipating this, no one will be deterred by the threat of punishment. Hence, although a large number of Nash equilibria (which allow incredible threats to be made) are possible in the CPR game with sanctioning, there is a unique subgame-perfect equilibrium in which all agents choose x_N and no agent sanctions any other.[5] Assuming common knowledge of the payoffs, this result applies for any finite repetition of the game.

It can be argued that infinite repetition of the one-shot CPR game is a more appropriate model for local commons. Then cooperative behavior can be sustained in subgame-perfect equilibrium. The limitations of this approach have already been discussed in the introduction. We adopt an alternative approach based on an evolutionary methodology in the sections to follow.

II. Evolutionary Dynamics

We now discard the assumption of purely self-interested behavior but retain the basic economic idea that individuals respond to differential payoffs by modifying their strategies. In keeping with the idea that behavioral patterns do not change instantaneously, we introduce some inertia in individuals' response to higher payoffs. The proportion of agents playing a particular strategy is subject to evolutionary pressure over time, with the population share of better performing strategies increasing relative to that of strategies earning lower payoffs. The payoffs to each strategy at any given point in time will depend, of course, on the prevailing population composition.

We wish to see whether there are outcomes other than that corresponding to the subgame-perfect equilibrium of the two-stage CPR game that may survive in the face of evolutionary pressure that weeds out strategies with low payoffs. To do this, we consider a simplified version of the model. Instead of a continuum of possible effort levels, we assume that there are only two, x_l and x_h, satisfying:

$$x_E \leq x_l < x_h \leq x_N. \tag{3}$$

It may be the case for example, that $x_l = x_E$ and $x_h = x_N$, so that agents can choose an effort level that is consistent with efficiency, or one that is consistent with the unique subgame-perfect equilibrium of the game considered in the previous section. Following common practice, players adopting high effort (and extraction) levels are referred to as *defectors*, since they impose as a result of their behavior a negative externality on the rest of the community. Those adopting low effort levels may be of two types: *enforcers* who sanction defectors, and *cooperators* who do not.[6] Note that this game, which has three strategies in the nor-

[5] David Hirshleifer and Eric Rasmusen (1989) show that cooperation can be sustained in subgame-perfect equilibrium in a finitely repeated prisoner's dilemma with sanctioning, but crucial to their result is the assumption that sanctioning takes the form of ostracism which is costless in the last period. A subgame-perfect cooperative equilibrium exists in a game with costly sanctioning if those who fail to sanction defectors can themselves be sanctioned, those who fail to sanction those who fail to sanction defectors can also be sanctioned, and so on in an infinite regress. There is little empirical support for this model. Some case studies (Cordell and McKean, 1992; Lawrence J. Taylor, 1987) suggest that the sanctioning of those who fail to sanction transgressors is unlikely, while in others (Acheson, 1988) such higher-level sanctioning is explicitly ruled out.

[6] The restriction to just three strategies is not necessary. For example, one could allow for players who defect but sanction other defectors, those who punish only cooperators, and those who sanction everyone. It can be shown using Theorem 4.1 in Somanathan (1995) that the introduction of these and similar strategies will not result in any substantive change in the conclusions of the paper. Nor will there be a qualitative change in the results if enforcers fail to identify defectors perfectly, provided the number of possible extraction levels remains finite and the probability of detecting deviations is sufficiently large for each extraction level.

mal form, and which is a restricted version of the CPR game of the previous section, has a unique subgame-perfect Nash equilibrium in which all players defect, and none are sanctioned.

Resource extraction (and sanctioning) occurs continuously over time, with each player adopting exactly one of the three pure strategies at any moment. Let s_1, s_2, and s_3 denote the proportion of players who are cooperators, defectors and enforcers, respectively.[7] Since cooperators and enforcers share the same extraction level, aggregate resource exploitation is given by

$$X = (1 - s_2)x_l n + s_2 x_h n.$$

Note that as a consequence of (3), the aggregate exploitation level satisfies

$$X_E \leq X \leq X_N$$

regardless of the population composition. Hence by Proposition 1, the average product must exceed the wage:

$$A(X) > w. \tag{4}$$

Given the population shares s_i at any point in time, it is assumed that each of the $s_3 n$ enforcers sanctions each of the $s_2 n$ defectors exactly once. The payoffs to each strategy type, given the population composition of strategies, are therefore:

$$\pi_1 = x_l(A(X) - w) \tag{5}$$

$$\pi_2 = x_h(A(X) - w) - s_3\delta n \tag{6}$$

$$\pi_3 = \pi_1 - s_2\gamma n. \tag{7}$$

Note that defectors' higher extraction levels give them higher payoffs than cooperators unless the damage they suffer from being sanctioned by enforcers is large enough to offset this advantage. Cooperators and enforcers do equally well if no defectors are present. Otherwise enforcers do worse than cooperators since they bear the costs of punishing defectors. Hence the enforcer strategy is weakly dominated by the cooperator strategy.

The payoff differentials will exert evolutionary pressure on the population composition, which we may expect to evolve in favor of those groups earning the highest payoffs. The simplest model of such movements in the population composition is that of *replicator dynamics* (Peter Taylor and Leo Jonker, 1978; Eric C. Zeeman, 1979), defined as

$$\dot{s}_i = s_i(\pi_i - \bar{\pi}), \qquad i = 1, 2, 3, \tag{8}$$

where $\bar{\pi} = \Sigma\, s_i\pi_i$ is the average payoff in the population as a whole. The rate of growth of the share of the population using a strategy is proportional to the amount by which that strategy's payoff exceeds the average payoff of the strategies in the population. These dynamics can be derived from models of individual learning behavior (John Gale et al., 1995; Jonas Björnerstedt and Jörgen Weibull, 1994), but we make no attempt to do so here. It is easily verified that the population shares s_i always sum to one and remain nonnegative under the replicator dynamics.[8]

From an evolutionary perspective, the main questions to be addressed are the following: what are the stable equilibrium points of the replicator dynamics, that is, which strategy or strategy combinations can survive in the long run? Is the unique subgame-perfect equilibrium in which all players defect the only possible long-run outcome? Or are there other strategies or strategy combinations characterized by restraint in resource extraction that can survive indefinitely?

To formulate and answer these questions precisely, we introduce the following terminology: a *state* of the dynamical system (8) is a vector of shares s_i. A *steady state* or

[7] Since n is finite, and only pure strategies are played, admissible values of s_i will be discrete. We abstract from this and allow the population shares to take on any nonnegative values which sum to unity. In most local commons, n ranges from a few tens to a few hundreds.

[8] In fact, using Theorem 4.1 in Somanathan (1995), the results of the paper can be shown to hold for all dynamics which satisfy *monotonicity* ($\dot{s}_i/s_i > (=)\ \dot{s}_j/s_j$ if $\pi_i > (=)\ \pi_j$), *Lipschitz continuity*, and *regularity* (see Larry Samuelson and Jianbo Zhang [1992] for definitions).

equilibrium point or simply *equilibrium* of the system is a state at which $\dot{s}_i = 0$ for all i. A *stable* equilibrium point is one such that for each neighborhood U of the point there exists a neighborhood U_1 of the point contained in U such that starting from a state in U_1 the state of the dynamical system will never leave U. An *asymptotically stable* equilibrium is one which is stable and has a neighborhood such that, starting in the neighborhood, the state of the system will converge to the equilibrium in the long run.

Since the population shares must always sum to unity, we may substitute $1 - s_1 - s_2$ for s_3 to obtain the following two-dimensional system:

$$\dot{s}_i = s_i(\pi_i - \bar{\pi}), \qquad i = 1, 2, \tag{9}$$

where $\bar{\pi} = s_1\pi_1 + s_2\pi_2 + (1 - s_1 - s_2)\pi_3$ is the average community-wide payoff. As a first step toward identifying stable states of the dynamics, consider the system's equilibrium points. A necessary and sufficient condition for a state to be an equilibrium point of the dynamics (9) is that all surviving strategies earn equal payoffs. The state consisting only of defectors is clearly an equilibrium point, as is any state in which no defectors are present. Let us refer to the steady state consisting only of defectors as the D-equilibrium, and the steady states in which no defectors are present as the C–E continuum. Clearly, a state in which all three types are present cannot be an equilibrium point since enforcers then do strictly worse than cooperators. A state consisting of only cooperators and defectors cannot be an equilibrium point either, since the latter get higher payoffs. There may exist an equilibrium point in which only defectors and enforcers are present but, as established in the result below, no such state can be stable.

PROPOSITION 2: *The D-equilibrium is asymptotically stable for all parameter values. In addition, if* $\delta n > (x_h - x_l)(A(nx_l) - w)$, *then the subset of stable points* **s** *in the C–E continuum constitute a nonempty interval* $\mathcal{S} = \{(s_1, s_2) \mid 0 \le s_1 < \bar{a}, s_2 = 0\}$ *where* $\bar{a}$ *is some positive number less than one. For each* **s** *in* $\mathcal{S}$, *there exists a neighborhood of* **s** *such that any trajectory originating in this neighborhood converges asymptotically to a point in* $\mathcal{S}$. *No other steady state is stable.*

According to Proposition 2, the D-equilibrium, which involves noncooperative resource exploitation and zero sanctioning, is locally stable for all parameter values. This is a very intuitive result since, when the proportion of enforcers in the population is sufficiently small, defectors perform strictly better than the other two groups. More interestingly, if the damages from sanctions are sufficiently high relative to the benefits derived from self-interested resource exploitation, there exists a set S of stable states which consist of a mix of cooperators and enforcers. Furthermore, given any state in $\mathcal{S}$, if the initial population composition is sufficiently close to this state, then the population will be driven (asymptotically) under evolutionary pressure into this set.[9] The intuition for this is that when there are sufficiently many enforcers, the proportion of defectors declines rapidly. With sufficiently few defectors, the payoff differential between cooperators and enforcers is small (it costs little to sanction very few people), so enforcers decline less rapidly. Consequently, defectors are eliminated before the enforcer share falls too much. The dynamics are illustrated in Figure 2.[10]

[9] It is known that imperfect equilibria can be dynamically stable. Examples have been given by Eric van Damme (1987), Samuelson and Zhang (1992), and Gale et al. (1995). Sethi (1996) finds imperfect equilibria to be evolutionarily stable if the set of players is augmented by a discriminating best-responder along the lines of Abhijit V. Banerjee and Weibull (1994). Proposition 2 provides an analytical explanation for the simulation results reported by Robert Axelrod (1986), since the CPR game with sanctioning is very similar in structure to the Norms Game considered there.

[10] Since the stable cooperative equilibria of Proposition 2 are not asymptotically stable, the reader may wonder whether the points in $\mathcal{S}$ will not be destabilized if defectors are reborn from time to time, which can result in the enforcer share falling gradually until the state drifts out of the basin of attraction of $\mathcal{S}$. If there is a (deterministic) drift in favor of *only* defectors, then the cooperative outcome is no longer stable. However, if the shares of all three strategies are subject to drift, then there is a range of parameter values in which the cooperative outcome is indeed stable. In other words, the main conclusions of Proposition 2 hold for many (though not all) deterministic perturbations of the replicator dynamics, not just for the replicator

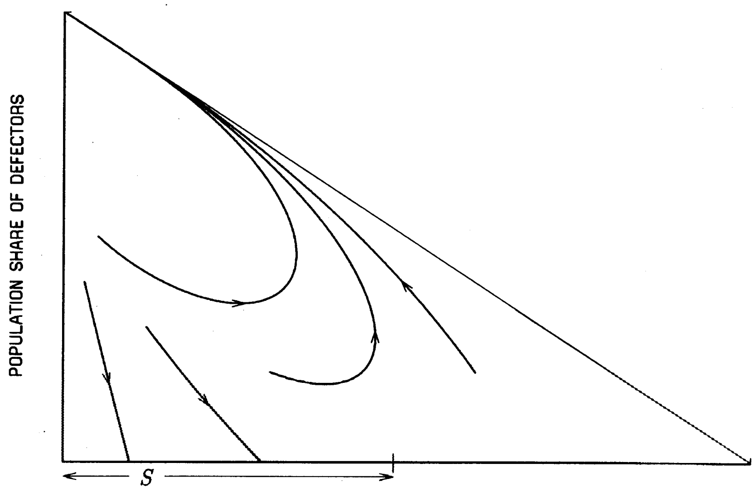

FIGURE 2. TWO CLASSES OF STABLE EQUILIBRIA

Hence there are two possible types of stable states towards which the system can evolve in the long run: an individualistic society of defectors, and a norm-guided one of cooperators and enforcers. Initial conditions, therefore, are crucial in determining the eventual outcome. The economic significance of this result is worth emphasizing. If C–E equilibria of the CPR game are understood as representing a departure from fully self-interested behavior in favor of norms of cooperation and decentralized enforcement, then the result establishes that there are conditions under which such a norm-guided population will be immune from invasion by a sufficiently small group of self-interested players. This provides a theoretical explanation for the overwhelming evidence both from the field and from experiments in favor of the existence of such norms in real-life finite-horizon CPR settings. It also provides some insight into the factors which can lead to a breakdown of such norms.

The size of the set of stable states and its corresponding basin of attraction will depend on the parameters of the model. It may be verified from the proof of Proposition 2 that the length of the set of stable states $\mathcal{S}$ is given by

$$1 - \frac{(x_h - x_l)(A(nx_l) - w)}{\delta n}$$

provided that the above expression remains nonnegative.[11] If the above expression is negative, no stable cooperative equilibrium exists. Hence a decline in δn reduces the size of the set of stable states, and sufficiently large declines in δn can cause it to vanish altogether. The intuition is clear: the payoff of defectors always exceeds that of cooperators and enforcers if the punishment they face is

dynamics alone. This is shown by means of numerical examples in the Appendix, which also discusses the effects of stochastic perturbations.

[11] Note that the length of the set of stable points $\mathcal{S}$ is independent of γ, the cost of sanctioning. The intuition for this is as follows. The stability of any point s in the C–E continuum depends on the payoffs to the three strategies at points close to s. Since the payoffs are continuous in the shares, the stability of s must depend only on the payoffs to the three strategies *at* s *itself*. But the enforcer payoff equals the cooperator payoff at s and the latter is independent of γ. The size of the basin of attraction of $\mathcal{S}$ does, however, depend on γ. This is perhaps most easily seen by noting that the edge of the basin of attraction on the s_2-axis must be an unstable steady state consisting of enforcers and defectors alone. The location of this steady state is given by the (smallest) value of s_2 that equalizes the payoffs of defectors and enforcers, which depends on γ.

sufficiently small. The smaller this punishment, the larger must be the enforcer population share for the low extraction level to be stable. Similarly, the size of the continuum of stable states is reduced when there is a *rise* in $A(nx_l) - w$, which is the average net return from harvesting the resource when all agents adopt low exploitation levels. The reasoning is straightforward: for any given intensity of sanctions, the greater the net return per unit of extractive effort, the greater will be the payoff of defectors relative to the other population types. A rise in $A(X)$ can come from two sources: an improvement in harvesting technology or a rise in the price of the resource.

Hence a norm of restraint can break down as a result of one of two independent factors: a decline in the intensity of social sanctions, or a rise in the net returns from resource extraction. Consider each of these in turn. Since sanctions often take the form of local penalties such as exclusion from cultural activities, the damage caused by such sanctions is liable to decline with the degree of cultural isolation of the community from the world at large. As the community becomes culturally integrated into a larger social entity, means of escaping local sanctions become available, and their psychological impact is prone to diminish. An increase in the net return from harvesting the resource, which gives rise to the same effect, can also come about through greater integration as superior harvesting technologies become available, and the local value of the resource rises with the prospect of sales in an external market. This effect might be mitigated by a rise in w, the opportunity cost of labor. Unfortunately, the effect of economic integration on wages has often lagged behind the other effects.

Notice that in this evolutionary framework, a temporary change in payoffs can bring about an irreversible change in behavior. Suppose, for example that economic integration increases the gains from defection, and this results in a move to the D-equilibrium from a cooperative one. Even if a subsequent rise in wages reverses the increase in the gains from defection, the stability of the D-equilibrium ensures that there will be no return to cooperative behavior. In a repeated-game optimizing model, on the other hand, there would be nothing that ruled out a return to a cooperative equilibrium that had temporarily ceased to exist because of a transitory shock to the parameters.

Population growth, represented by an increase in n, has several effects. First, note that δ, x_l, and w, so far taken as exogenous, are all likely to vary with n. Increases in group size, which bring with them anonymity, will probably diminish the force of social sanctions. We can therefore write $\delta = \delta(n)$, with $\delta(n)n$ decreasing in n in the relevant range. This will tend to shrink $\mathcal{S}$ as n increases. The increase in population, by raising demand, can raise the price of the resource and thus the function $A(\cdot)$. It can also, through an increase in labor supply, reduce w. These effects also tend to shrink S. If the norm for individual labor input x_l is derived from a norm for the aggregate harvest, then x_l will be reduced so as to leave nx_l unchanged. This increases $(x_h - x_l)$ which also reduces the length of $\mathcal{S}$. However, if x_l does not vary with n, then $A(nx_l)$ will fall, which tends to reduce the gains from defection, albeit at the cost of making the norm itself less efficient. To summarize, population growth will probably, though not necessarily, reduce the length of the set of stable points.

The main prediction of the model, as developed thus far, is that a gradual change in the model parameters, arising for instance from increasing external contact, will have negligible effects on the behavior of a community *as long as the population composition remains within the basin of attraction of the set of stable states*. If this set of stable states continues to shrink over time, there may arise a point at which the population composition falls outside it's basin of attraction: in this case one may expect to see a rapid and catastrophic breakdown of the norm as the population starts to move in the direction of the defector equilibrium. The effect of this shift on the dynamics of the resource stock is explored in the following section.

III. Resource-Stock Dynamics

The model has hitherto been based on the assumption of an exogenously given resource stock, which may be considered reasonable only if the aggregate harvest is negligible in comparison with the natural rate of replenishment of the resource. Despite the restrictive nature of this

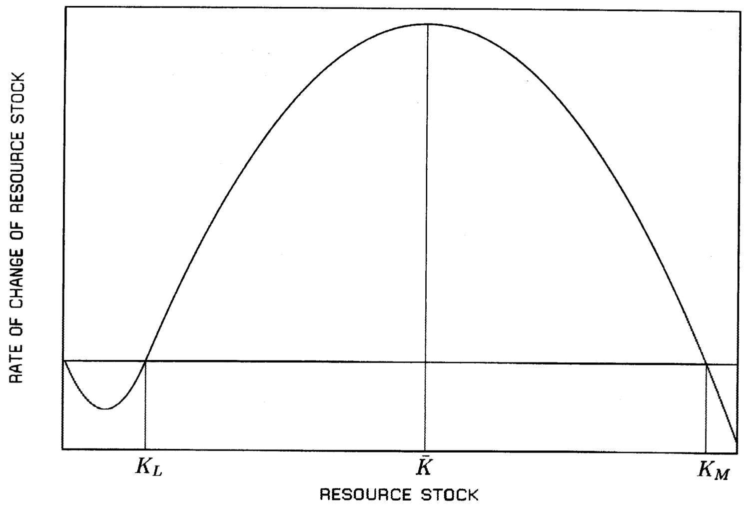

FIGURE 3. STOCK REGENERATION IN THE ABSENCE OF HARVESTING ($\dot{K} = G(K)$)

assumption, the results of Section II do shed some light on the dynamics of the resource stock. It is clear, for instance, that each stable state of the system considered in the previous section corresponds to a given equilibrium rate of resource extraction. If this rate is below the maximum sustainable yield of the resource, then the extraction can be sustained indefinitely. Since the stable equilibrium consisting exclusively of defectors gives rise to a rate of extraction $f(nx_h)$ which exceeds the extraction rate $f(nx_l)$ in the C–E continuum, it is clear that a resource stock able to sustain a population of the former type will be able to sustain a population of the latter type. The converse need not hold: a resource able to sustain a population of the C–E type indefinitely may be fully depleted if the population evolved towards the defectors. This possibility can be fully explored only if the resource-stock dynamics are explicitly taken into account.

In order to endogenize K, recall the original harvest function $H(X, K)$ introduced in Section I. This corresponds to a family of curves such as that depicted in Figure 1, with each curve corresponding to a different stock K. In addition to its dependence on the harvest, the evolution of the resource stock will depend on its own natural rate of replenishment, which we represent by the differentiable function $G(K)$. There is a finite *carrying capacity* K_M of the resource stock so that $G(K) < 0$ for $K > K_M$ and $G(K_M) = 0$. Let $K_L > 0$ be the minimum viable stock, so that $G(K) > 0$ for $K_L < K < K_M$, and $G(K) < 0$ for $0 < K < K_L$. The minimum viable stock is the level below which the resource cannot recover by natural reproduction even in the absence of harvesting. For some resources this may be close to zero, so that cessation of harvesting will lead to stock recovery even when the resource is close to extinction but this situation is far from general.[12] Finally, assume G has a unique maximum at some $\bar{K}$. This is the standard specification used to characterize the dynamics of renewable resources. A function satisfying these properties is depicted in Figure 3.

The evolution of the resource stock, taking account of harvesting, is then given by

$$\dot{K} = G(K) - H(X, K). \tag{10}$$

Note that harvesting is worthwhile only if $H_X(0, K) \geq w$. Let $K_{\min}$ be the minimum K for which this is true ($K_{\min} \geq 0$). Letting $x_E(K)$ denote the (statically) socially optimal

[12] Colin W. Clark (1990 p. 20) discusses the case of the Antarctic blue whale the recovery of which, despite almost three decades of a total ban on harvesting, remains in doubt.

average effort level and $x_N(K)$ denote the Nash-equilibrium effort level corresponding to the resource stock K, it is clear that $x_E(K) = x_N(K) = 0$, for $K \leq K_{\min}$. Now we can apply Proposition 1 to every $K > K_{\min}$. For each such K, there corresponds a unique $x_E(K) > 0$ representing the (statically) efficient symmetric level of exploitation, and a unique $x_N(K) > x_E(K)$ representing the unique noncooperative equilibrium level of exploitation. It may be shown that these are increasing in K. We use this as motivation for the assumptions that

$$x_l(K) = x_h(K) = 0, \qquad \text{for } K \leq K_{\min},$$

and that

$$x_l(K),\, x_h(K) \text{ are positive and increasing for } K > K_{\min}.$$

We are now allowing the noncooperative extraction level x_h and the norm-guided extraction level x_l to vary with the resource stock. Norms which permit resource use to increase with the stock are commonly observed. For example, in Kumaun, women lopping oak leaves for fodder are expected to leave a certain proportion of the leaves on each branch. In the Maine lobster fisheries, only lobsters within a particular size class can be legally harvested, with the enforcement of the law being left to norms (Acheson, 1988).

As before, we suppose that (3) holds for all $K > K_{\min}$. Assume further that $x_l(K)$ and $x_h(K)$ are continuous.[13] Since extraction levels for each type of player now depend on K, so does aggregate extraction X:

$$X(s_2, K) = (1 - s_2)x_l(K)n + s_2 x_h(K)n.$$

The average product now depends both on aggregate extraction effort and the resource stock

$$A(X, K) = \frac{H(X, K)}{X}.$$

The payoffs of the three player types remain the same but we can rewrite them to reflect their dependence on K:

$$\pi_1 = x_l(K)(A(X, K) - w)$$

$$\pi_2 = x_h(K)(A(X, K) - w) - (1 - s_1 - s_2)\delta n$$

$$\pi_3 = x_l(K)(A(X, K) - w) - s_2\gamma n.$$

Recalling the replicator dynamics from Section II, we have

$$(11) \qquad \dot{s}_i = s_i(\pi_i - \bar{\pi}), \qquad i = 1, 2,$$

where $\bar{\pi} = s_1\pi_1 + s_2\pi_2 + (1 - s_1 - s_2)\pi_3$ as before. Equation (10) in conjunction with the two equations in (11) constitute a three-dimensional dynamical system in the variables s_1, s_2, and K.

We are interested in the stable equilibrium points of this system, that is, the types of behavior and levels of resource stock that we may expect to see in the long run. First consider the aggregate effort levels X and resource stocks K for which $\dot{K} = 0$. For any $K_L < K < K_M$, $G(K) > 0$, so by putting in enough effort, the harvest can be raised high enough that it equals the rate of replenishment, thus causing $\dot{K}$ to equal zero.[14] For any such K, let $\hat{X}(K)$ be the level of effort which will make $\dot{K} = 0$. It is straightforward to verify that $\hat{X}(K)$ is decreasing in the range $(\bar{K}, K_M)$ and attains its maximum before $\bar{K}$. Figure 4 depicts a typical $\dot{K} = 0$ isocline, as well as the increasing functions $X_h(K) = nx_h(K)$ and $X_l(K) = nx_l(K)$.

Observe that aggregate extraction $X(s_2, K)$ must lie between the curves $X_l(K)$ and $X_h(K)$ and that $\dot{K} > 0$ below $\hat{X}(K)$, and less than zero above it. Next, note that the population composition is invariant on $X_l(K)$ and $X_h(K)$ (because these correspond to a mix of cooperators and enforcers, and to only defectors, respectively). So the intersections of these curves with $\hat{X}(K)$ correspond to steady states of the system. By Proposition 2, any equilib-

[13] As in Section II, a special case of this formulation occurs if $x_l(K) = x_E(K)$, and $x_h(K) = x_N(K)$.

[14] We assume this property of H, without which the resource would be inexhaustible.

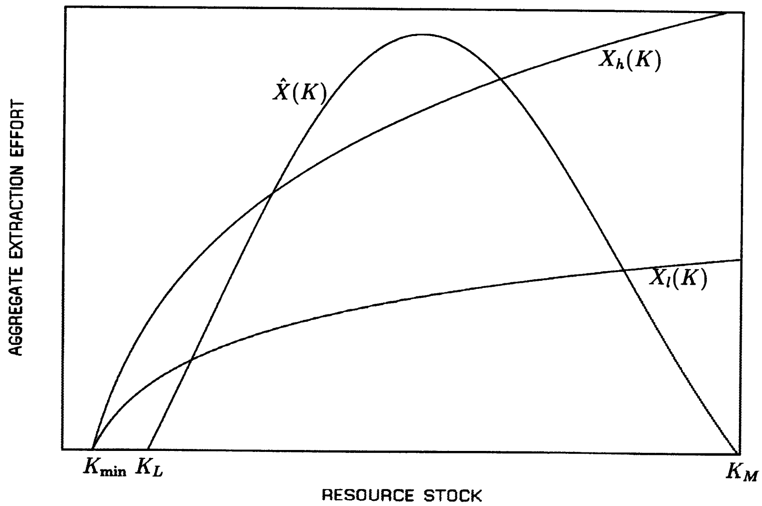

FIGURE 4. STEADY STATES WITH EVOLUTIONARY AND RESOURCE STOCK DYNAMICS

rium in which defectors coexist with enforcers is unstable, so the intersections correspond to the only possible stable states. The stability properties of states consisting only of defectors are easily characterized (primes refer to first derivatives, and a D-equilibrium is a point at which $s_2 = 1$ and $\dot{K} = 0$).

PROPOSITION 3: *Any D-equilibrium is asymptotically stable if $X_h'(K) > \hat{X}'(K)$ and unstable if $X_h'(K) < \hat{X}'(K)$ at the corresponding resource stock.*

In other words any intersection of the $X_h(K)$ curve with the $\hat{X}(K)$ isocline yields a stable equilibrium if the slope of the former curve exceeds that of the latter. There may, of course, be no such intersection if the returns to harvesting are sufficiently high even at low levels of the resource stock. ($X_h(K)$ may lie entirely above $\hat{X}(K)$.) This issue is reconsidered below. With regard to the existence of stable equilibria with norm-guided players, the following holds.

PROPOSITION 4: *If there exists a stable D-equilibrium then there exists a continuum of C–E equilibria with a higher resource stock. Furthermore, if $\delta n > (x_h - x_l)(A - w)$ and $\hat{X}(K)' < 0$ at the resource stock corresponding to the C–E continuum, then there exists a nonempty subset $\mathcal{S}$ of this continuum such that every point* **s** *in $\mathcal{S}$ corresponds to a stable state in the* $(\mathbf{s}, K)$ *dynamics* (10)–(11).

The proposition states that if there exists a stable equilibrium in which only defectors survive, then there exists a continuum of equilibria with a higher resource stock in which only cooperators and enforcers survive. If this continuum corresponds to a resource stock at which the isocline $\hat{X}(K)$ is downward sloping, and provided the damages from sanctions are sufficiently high, then there exists a subset of this continuum composed of stable equilibria. Note that there may exist stable equilibria with a positive resource stock consisting exclusively of norm-guided players when there exist none consisting exclusively of self-interested types (that is when $X_h(K)$ lies entirely above $\hat{X}(K)$).

The significance of Propositions 3 and 4 can best be appreciated by examining the (nonexhaustive) list of possibilities depicted in Figure 5. In Figures 5A and 5B, the marginal returns to effort begin to exceed the wage at levels of the resource stock below the minimum viable stock ($K_{\min} < K_L$). Under these conditions, if there exists a stable D-equilibrium with a positive resource stock, as in Figure 5A, there must exist a continuum of equilibria at a *higher* resource stock in which defectors are extinct. If δ is sufficiently large there will be

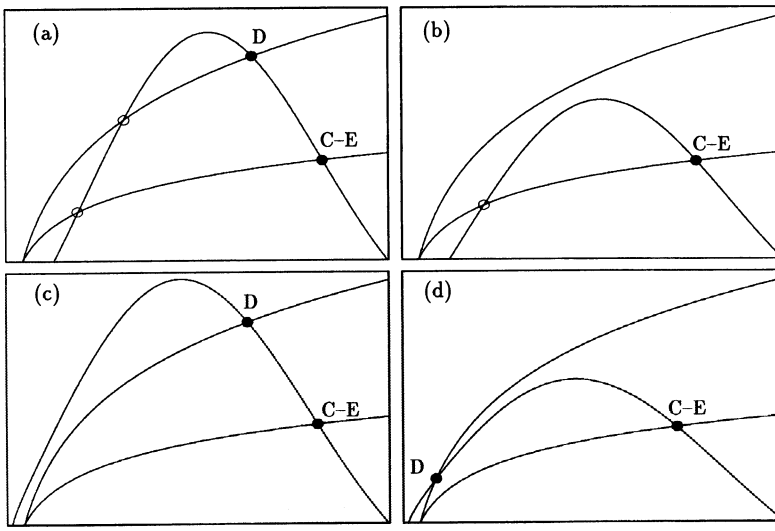

FIGURE 5. SOME POSSIBLE CONFIGURATIONS OF EQUILIBRIA (● STABLE, ○ UNSTABLE)

a subset of this continuum which consists of dynamically stable points. On the other hand, for some parameter values, there may exist a continuum of stable equilibria in which only norm-guided players survive but none in which they are absent. This situation is depicted in Figure 5B. Under these conditions a positive resource stock can only be sustained to the extent that the social norms of restraint and enforcement remain intact. A breakdown of such norms will lead not simply to an increase in extraction effort and a disappearance of sanctioning but to the complete extinction of the resource stock.

Figures 5C and 5D depict situations in which the marginal returns to effort start to exceed the wage only after the minimum viable stock has been exceeded ($K_{\min} > K_L$). Extinction resulting from a breakdown of social norms is impossible in these cases. In accordance with Proposition 3, there must exist a stable equilibrium with a homogeneous population of self-interested players. Furthermore, if the additional requirements of Proposition 4 are satisfied, there will exist a continuum of stable points in which only norm-guided players persist. Note that the latter class of stable equilibria *always* entail a higher equilibrium value of the resource stock.

We are now ready to examine the effects of changes in the parameters on the equilibrium levels of the resource stock. Suppose we are near a cooperative equilibrium (a point in $\mathcal{S}$). A fall in the damage inflicted by sanctions or a rise in the net return to resource extraction will shrink the set $\mathcal{S}$ of stable points. As long as the current state remains within the basin of attraction of $\mathcal{S}$, there will be no change in the share of each behavior and so no change in the stock. But if the parameter values continue to change there must come a time when $\mathcal{S}$ shrinks so much that the current state leaves its basin of attraction. The resulting rise in the proportion of defectors will now bring about a fall in the resource stock. As the proportion of defectors continues to rise (which will happen with increasing rapidity as it gets more and more costly to adhere to the norm) the stock will fall further until a stable D-equilibrium is reached at a lower level of K. Notice that the equilibrium resource stock can fall sharply in response to an imperceptibly small parameter change as points near the current state lose stability, which distinguishes this scenario from what would happen if all agents were using the one-shot Nash equilibrium extraction level $x_N(K)$. In that case, a rise in the resource price would shift the $X_h(K)$ curve up and the equilibrium stock would fall smoothly.

In Figures 4 and 5, in (K, X)-space, there will be a movement up from a stable C–E equilibrium. Once above the $\hat{X}(K)$ curve, $\dot{K}$ becomes negative. In cases like those depicted in Figures 5A and 5C, in which decreasing returns to effort are sufficiently pronounced, a new equilibrium at a *lower* resource stock and a *higher* extraction effort is reached. If agents are sufficiently patient, this is a welfare reducing change for the following reason.[15] For any given extraction effort a lower resource stock corresponds to a diminished flow of output (a downward shift of the production function depicted in Figure 1). Similarly, for any given resource stock the effort level x_l leads to a higher per capita harvest than the effort level x_h on account of (3). Hence movements to the north-west of an equilibrium point are always welfare reducing. Cases such as those depicted in Figures 5A and 5C may describe some inshore fisheries in which decreasing returns to effort for a given resource stock are quite pronounced. In the case of forest resources, on the other hand, noncooperative effort levels may be consistent only with a much lower resource stock or complete extinction, as in Figures 5B or 5D. Here a breakdown of a cooperative norm, if it does not engender resource extinction, may lead to a highly depleted long-run stock with *lower* effort levels as in 5D. In this case welfare can be shown to have declined in the special case $x_l = x_E$, where cooperative extraction levels are statically efficient. This follows from the fact that at the lower stock, x_l yields a higher per capita harvest than x_h, and movements upward along the $x_l(K)$ locus are welfare increasing, since they correspond to upward shifts of the production function depicted in Figure 1. Hence, provided that cooperative extraction levels are sufficiently close to statically efficient levels, the breakdown of a cooperative norm can quite generally be expected to lead not only to a decline in the long-run resource stock but also to a corresponding decline in welfare.

A decline in the punitive effect of social sanctions, which could trigger the breakdown of a cooperative norm, can result from the incursions of outsiders immune to local sanctions. This will result in a rise in the share of defectors with the consequences detailed above. Historically, such an opening of access to local commons has been an important factor in resource depletion. This has stemmed from the failure of states to recognize the informal management systems operating and to take them into account when framing regulations and modifying property rights. To illustrate, consider the case of the forests of Kumaun and Garhwal in Northern India, discussed in detail in Somanathan (1991).

Prior to the British conquest in 1815, virtually the entire forest area was implicitly designated the exclusive common property of one village or another, with clearly demarcated boundaries recognized by members of adjacent villages. Intravillage patterns of use were regulated by norms, sometimes backed up by strict rules enforced by village councils. Between 1911 and 1917 vast areas of forest were taken over by the colonial government with a view to managing the extraction of timber from pines for railway construction. A common property regime was thereby transformed into a state property regime. The resulting restrictions on the use of forest resources by its inhabitants played havoc with the villagers' customary patterns of use and led to massive popular protests and huge incendiary fires. In 1925 an alarmed government removed virtually every restriction on the use of the less commercially valuable oak forests by "all *bona fide* residents of Kumaun," thus transforming a state property regime to one of open access. Not only were the former village boundaries rendered irrelevant, those for whom it was expedient to violate custom could now do so with the law on their side. The five-year period following this government order led to a rapid and unprecedented rate of deforestation as villagers suddenly immune from the restraining effects of custom and local sanction began encroaching into areas previously

[15] If agents are not sufficiently patient a movement to an equilibrium with lower steady-state welfare could be compensated for by an increase in welfare during the transition. The welfare effects of a transition from one state to another cannot be properly analyzed without an explicit intertemporal utility function in this case.

considered to be the property of others. The conservator of forests, E. A. Smythies, was led to observe in 1931 that "the oak is melting away in Kumaun like an iceberg on the equator."[16] The case of Kumaun is a striking illustration of both the potential viability of common property regimes when sustained by social norms, and the dangers of making alterations in existing property regimes in such a manner as to undermine the effectiveness of local sanctions.

IV. Conclusions

The model of common-property resource use developed here is characterized by decentralized exploitation and costly sanctions, and has the property that norms of restraint and punishment can be stable against invasion by narrowly self-interested players. The results provide a theoretical explanation of the overwhelming field and laboratory evidence describing frequent cooperation and sanctioning behavior in small groups. In addition, the model predicts persistence of behavior patterns. A temporary parameter change that destabilizes a cooperative equilibrium will have permanent effects on equilibrium behavior. Once trajectories enter the basin of attraction of the defector equilibrium, which is stable for all parameter values, cooperative behavior will be extremely difficult to recover. This prediction is consistent with the historical record and cannot be derived on the basis of the standard infinite-horizon optimizing framework. Moreover, the results of Section III apply to any social situation characterized by negative externalities and help explain Putnam's (1993) description of the deep historical roots of civic behavior.

Although the focus has been on fully decentralized enforcement, the model can be generalized in a straightforward manner to account for costly centralized monitoring funded by voluntary contributions.[17] A critical assumption in the model developed here is that there is no fixed cost of monitoring, independent of the degree of enforcement activity undertaken, which enforcers are obliged to pay. If this assumption were to be relaxed, effective enforcement would require some form of specialized monitoring, for instance through the use of paid guards. In the absence of a coercive state, such payments would have to come from voluntary contributions, which again opens the door to the free-rider problem. In this case, defectors may be identified with non-contributors, cooperators with contributors, and enforcers with contributors who sanction noncontributors. Sanctioning of those who overexploit the resource (as opposed to the sanctioning of noncontributors) may now be centralized. The scenario is complicated by the fact that the payoffs of those who are sanctioned for overexploiting the resource will depend on the population composition indirectly, through its effect on aggregate contributions and hence on the scope for centralized monitoring activity. Nevertheless, provided that sanctioning of noncontributors is sufficiently damaging and the monitoring of contributions is not itself costly, there will be a continuum of stable states in which resource exploitation is effectively supervised by a centralized monitoring authority funded by voluntary contributions. This corresponds to the case of the watchmen of Kumaun (Somanathan, 1991), whose salaries are paid out of a voluntary fund, and whose vigilance increases in direct proportion to their remuneration. Hence the framework developed here is capable of explaining restraint in resource use even when fixed costs of monitoring cause the monitoring mechanism to be centralized.

Equilibria characterized by cooperative norms may, however, be rendered unstable for a number of reasons. In the model considered here, this could arise as a conse-

[16] Quoted in Somanathan (1991). For other examples see Ostrom (1990 Chapter 5) on the inshore fisheries of Newfoundland and the communal forests of Nepal.

[17] We thank Martin Weitzman for bringing to our attention the importance of considering this case. The efficiency implications of centralized enforcement in CPR settings have been investigated by Franz Weissing and Ostrom (1993), who show that the degree of noncooperative behavior may rise or fall relative to the decentralized case. In the closely related context of collective action by interest groups, Jonathan Bendor and Dilip Mookherjee (1987) show that with imperfect monitoring, centralized enforcement can induce efficient outcomes in circumstances where decentralized enforcement fails. Both papers use rational-choice game-theoretic models.

quence of a rise in the price of the resource, or a diminution of the damage that sanctions such as cultural isolation entail. Alternatively, a breakdown could result from the erosion of common-property rights as state property or open-access regimes are imposed by a centralized authority, allowing outsiders immune from local sanctions to gain access. The implications of a loss of stability depend crucially on whether or not the marginal returns to effort exceed the opportunity cost of labor at magnitudes of the resource stock that are below the minimum viable level. If they do, there may arise a situation in which no stable equilibria with a positive resource stock and narrowly self-interested extraction levels exists. A breakdown of norms under such circumstances ensures resource extinction. This too is consistent with the historical record and suggests that governments should pay careful attention to existing norms of use when framing regulations and changes in property rights.

This analysis of the effects of exogenous changes distinguishes the theory of cooperative behavior presented here from theories based on self-interested agents interacting over an infinite horizon. The integration of the dynamics of the resource stock with the evolutionary dynamics and the consequent implications of exogenous changes for stock levels also distinguishes our model from previous work in evolutionary game theory. The model enables one to trace the effects of a change in observable parameters through to a change in behavior to a change in an observable stock, while previous work has been conducted entirely in terms of behaviors which might be difficult to observe. By spelling out the implications of changes in observable parameters for an observable stock, the theory is made empirically operational.

One limitation of the analysis is that it does not allow for institutional change. When the breakdown of norm-guided decentralized sanctioning threatens the viability of a vital resource, it is possible for more centralized enforcement mechanisms in the form of explicit laws, policing, and institutionalized punishment to evolve at the local level in response to the threat. This possibility, though empirically important, is outside the scope of the present investigation and must remain a topic for future research. It may be added however, that even when such new governance mechanisms play an important role in restraining resource extraction, the existence of norms enable them to function more smoothly and cost effectively.

The use of evolutionary dynamics here is an example of a model of the pursuit of individual payoffs by economic agents that does not assume optimizing behavior. It illustrates the possibility that this weakening of the standard postulate can lead to outcomes ruled out by the more usual methodology. The evolutionary approach has also enabled us to develop a theory that is consistent with the anthropological literature on common-property regimes, without discarding what we see as the centerpiece of economic reasoning: the tendency of human behavior to adjust in response to persistent differentials in material incentives.

APPENDIX

PROOF OF PROPOSITION 2:

The Jacobian of the two-dimensional system (9) is given by **J** below.

$$\mathbf{J} = \begin{pmatrix} \pi_1 - \bar{\pi} + s_1 \dfrac{\partial(\pi_1 - \bar{\pi})}{\partial s_1} & s_1 \dfrac{\partial(\pi_1 - \bar{\pi})}{\partial s_2} \\ s_2 \dfrac{\partial(\pi_2 - \bar{\pi})}{\partial s_1} & \pi_2 - \bar{\pi} + s_2 \dfrac{\partial(\pi_2 - \bar{\pi})}{\partial s_2} \end{pmatrix}$$

Now consider each of the three equilibrium types identified in Proposition 2.

Claim 1.—Stability of the D-equilibrium at (0, 1). In this case $\bar{\pi} = \pi_2$ and $X = nx_h$. The elements of the Jacobian (after some manipulation and simplification) are

$$J_{11} = -(x_h - x_l)(A(nx_h) - w) < 0$$

$$J_{12} = 0$$

$$J_{21} = -\gamma n < 0$$

$$J_{22} = -(x_h - x_l)(A(nx_h) - w) - \gamma n < 0.$$

The Jacobian therefore has a positive determinant and negative trace, conditions which are necessary and sufficient for local asymptotic stability.

Claim 2.—Stability of the continuum $\mathcal{S}$ of equilibria $(a, 0)$, $a \in [0, 1]$. Here $s_2 = 0$ and $\bar{\pi} = \pi_1 = \pi_3$. Hence $J_{11} = J_{21} = 0$, implying that the determinant of the Jacobian is zero. This means that one eigenvalue is zero and the other equals the trace of the Jacobian. The second eigenvalue therefore equals J_{22} which in this case equals $\pi_2 - \pi_1$. It is positive if

$$x_l(A(nx_l) - w) < x_h(A(nx_l) - w) - \delta n(1 - s_1)$$

which leads to

$$(1 - s_1) < \frac{(x_h - x_l)(A(nx_l) - w)}{\delta n}.$$

The right-hand side of this equation is positive due to (4). It is greater than 1 if $\delta n < (x_h - x_l)(A(nx_l) - w)$. In this case the system is unstable at $(a, 0)$.

Now suppose $\delta n > (x_h - x_l)(A(nx_l) - w)$. Then there exists $\bar{a} \in (0, 1)$ such that $(1 - \bar{a})\delta n = (x_h - x_l)(A(nx_l) - w)$. We now show that the system is locally stable at $(a, 0)$ for all $a < \bar{a}$. Let U be a neighborhood of $(a, 0)$. Now

$$\dot{s}_1 = s_2 s_1[(1 - s_1 - s_2)\gamma n + (1 - s_1 - s_2)\delta n - (x_h - x_l)(A(X) - w)],$$

and

$$\dot{s}_2 = s_2\{-(1 - s_2)[(1 - s_1 - s_2)\delta n - (x_h - x_l)(A(X) - w)] + (1 - s_1 - s_2)\gamma n s_2\}.$$

Choose $\varepsilon > 0$ such that the rectangle $R \equiv \{(s_1, s_2) : a - \varepsilon/2 \leq s_1 \leq a + \varepsilon/2, 0 < s_2 \leq \varepsilon/2\}$ is contained in U and $a + \varepsilon < \bar{a}$. So within the rectangle $s_1 + s_2 < \bar{a}$ and so $(1 - s_1 - s_2)\delta n - (x_h - x_l)(A(X) - w) > 0$, hence $\dot{s}_1 > 0$. Note also that $\dot{s}_1/s_2$ is bounded above by (say) M.

Also, ensure that ε is chosen to be small enough that for all (s_1, s_2) in R, $\dot{s}_2/s_2$ is bounded away from zero, by (say) L. So within R, $0 < \dot{s}_1/-\dot{s}_2 < M/L$. Note that $\dot{s}_1/-\dot{s}_2$ is the tangent of the angle the vector field at (s_1, s_2) makes with the s_2-axis. Now consider any triangle Δ with vertices $(x, 0)$, $(x, \hat{s}_2)$, $(x + \hat{s}_2(M/L), 0)$ that contains a and is contained in the closure of R. The vector field at any point in the interior of the triangle now points towards the edge of the triangle that lies on the s_1-axis. Since all points on this edge are invariant, the flow from any point in the interior cannot leave the triangle and so cannot leave U. Hence $(a, 0)$ is locally stable.

From the above, we know that for every point $(a, 0)$ in the set $\mathcal{S}$, there is a neighborhood U that is an invariant set. Choose one such neighborhood for each point in $\mathcal{S}$ and denote the union of these neighborhoods by V. Then V is itself an invariant set. Since we are dealing with a planar flow, all trajectories originating in V must converge either to an equilibrium point or to a limit cycle (by the Poincare-Bendixson theorem). We can rule out limit cycles since each limit cycle must enclose an equilibrium point (John Guckenheimer and Philip Holmes, 1983 p. 51), and there are

$$\mathbf{J} = \begin{pmatrix} s_1 \dfrac{\partial(\pi_1 - \bar{\pi})}{\partial s_1} + \pi_1 - \bar{\pi} & s_1 \dfrac{\partial(\pi_1 - \bar{\pi})}{\partial s_2} & s_1 \dfrac{\partial(\pi_1 - \bar{\pi})}{\partial K} \\ s_2 \dfrac{\partial(\pi_2 - \bar{\pi})}{\partial s_1} & s_2 \dfrac{\partial(\pi_2 - \bar{\pi})}{\partial s_2} + \pi_2 - \bar{\pi} & s_2 \dfrac{\partial(\pi_2 - \bar{\pi})}{\partial K} \\ -H_X \dfrac{\partial X}{\partial s_1} & -H_X \dfrac{\partial X}{\partial s_2} & G' - H_X \dfrac{\partial X}{\partial K} - H_K \end{pmatrix}$$

no equilibrium points in the interior of the 2-simplex. Hence all trajectories originating in V must converge to an equilibrium point in V. The only such points are also in $\mathcal{S}$.

Claim 3.—Instability of remaining equilibria. The only remaining candidate for a stable equilibrium is one consisting exclusively of defectors and enforcers. Here $s_1 = 0$ and $\bar{\pi} = \pi_2 = \pi_3 < \pi_1$ since $\pi_3 < \pi_1$ whenever $s_2 > 0$. Hence $J_{11} = \pi_1 - \bar{\pi} > 0$. Also, since $s_1 = 0$, we have $J_{12} = 0$. If $J_{22} \geq 0$ then the trace of the Jacobian is positive. If, on the other hand, $J_{22} < 0$ then the determinant of the Jacobian is negative. In either case, the equilibrium cannot be stable.

PROOF OF PROPOSITION 3:

The Jacobian matrix for the system (11) is given by **J** above.

At any D-equilibrium $(0, 1, K^*)$, evaluation of the above expressions reveals that $J_{12} = J_{13} = J_{23} = 0$. Hence the Jacobian is a lower triangular matrix, so that its eigenvalues are equal to its three diagonal elements. These are given by

$$J_{11} = -(x_h - x_l)(A(X^*) - w) < 0$$

$$J_{22} = -(x_h - x_l)(A(X^*) - w) - \gamma n < 0$$

$$J_{33} = G' - H_X \frac{\partial X_h}{\partial K} - H_K,$$

where $X^* = X(1, K^*)$ is the equilibrium extraction effort and the signs follow from the fact that the average product exceeds the wage regardless of the population composition. Hence all eigenvalues of the system are real, and at least two are strictly negative. The sign of the third is also negative if

$$X_h' > \frac{G' - H_K}{H_X}$$

and positive if the sign of the above inequality is reversed. But the right-hand side of the inequality is just $\hat{X}'$ (applying the implicit function theorem to $G(K) - H(X, K)$).

PROOF OF PROPOSITION 4:

Existence is obvious from Figures 4 and 5. To prove stability note that by continuity of x_l, x_h, and A, $\delta n - (x_h(K) - x_l(K))(A(nx_l(K), K) - w)$ is bounded away from zero in some neighborhood $(K^{**} - k, K^{**} + k)$ of the equilibrium resource stock K^{**}. Hence there exists $\bar{a}$ such that $(1 - a)\,\delta n - (x_h(K) - x_l(K))(A(nx_l(K), K) - w) > 0$ for all $a < \bar{a}$ in this neighborhood of K^{**}.

Now follow the proof of Claim 2 of Proposition 2, with the addition that ε is chosen small enough that $X(s_1, s_2, K) < \hat{X}(K^{**} - k)$ for all (s_1, s_2) in R and K in $(K^{**} - k, K^{**} + k)$. If we start at a point of $\Delta \times (K^{**} - k, K^{**} + k)$ at which $K > K^{**}$ then $\dot{K} < 0$, so the flow points back into this neighborhood as before. If we start at a point at which $K < K^{**}$, it may be the case that $\dot{K} < 0$ even now. However X must be decreasing and hence (since $\hat{X}(K)$ is downward-sloping), K must stop decreasing before it can leave

TABLE A1—EIGENVALUES OF THE JACOBEAN FOR VARIOUS MUTATION RATES

Mutation rate	Equilibrium shares				Eigenvalues	
ρ	s_1	s_2	s_3	Type	λ_1	λ_2
0.005	0.563	0.001	0.436	C–E	−0.006	−1.920
	0.009	0.989	0.001	D	$-0.704 + 0.739i$	$-0.704 - 0.739i$
0.010	0.562	0.002	0.436	C–E	−0.012	−1.913
	0.019	0.978	0.003	D	$-0.730 + 1.068i$	$-0.730 - 1.068i$
0.015	0.562	0.003	0.436	C–E	−0.016	−1.907
	0.028	0.967	0.004	D	$-0.754 + 1.243i$	$-0.754 - 1.243i$
0.020	0.561	0.004	0.435	C–E	−0.020	−1.901
	0.038	0.956	0.006	D	$-0.777 + 1.332i$	$-0.777 - 1.332i$
0.025	0.560	0.004	0.435	C–E	−0.024	−1.896
	0.047	0.945	0.008	D	$-0.800 + 1.354i$	$-0.800 - 1.354i$
0.030	0.560	0.005	0.435	C–E	−0.027	−1.892
	0.057	0.934	0.009	D	$-0.821 + 1.317i$	$-0.821 - 1.317i$
0.035	0.559	0.006	0.435	C–E	−0.029	−1.888
	0.067	0.922	0.011	D	$-0.841 + 1.217i$	$-0.841 - 1.217i$
0.040	0.559	0.007	0.434	C–E	−0.030	−1.884
	0.077	0.911	0.013	D	$-0.861 + 1.041i$	$-0.861 - 1.041i$
0.045	0.558	0.008	0.434	C–E	−0.031	−1.882
	0.087	0.899	0.015	D	$-0.879 + 0.739i$	$-0.879 - 0.739i$
0.050	0.557	0.009	0.434	C–E	−0.031	−1.880
	0.097	0.887	0.016	D	−1.250	−0.544

$(K^{**} - k, K^{**} + k)$. Hence the flow can never leave the neighborhood and the equilibrium is stable.

Perturbations of the Dynamics

We wish to consider the robustness of Proposition 2 with respect to perturbations of the dynamics. First, we demonstrate robustness to deterministic perturbations.

Consider the perturbed replicator dynamics (Gale et al., 1993) which allow for a perpetual rate of drift from one strategy to another:

$$\text{(A1)} \quad \dot{s}_i = (1-\rho)s_i(\pi_i - \bar{\pi}) + \rho(\theta_i - s_i),$$

$$i = 1, 2.$$

The interpretation of the above system is the following. At any point in time, a fraction $1 - \rho$ of the population adapts according to the standard replicator equations used in the text, while the remainder ρ are replaced by agents who follow pure strategy i with frequency θ_i. To show that norm-guided cooperation and enforcement can be stable even in the presence of such drift, we have computed equilibria of the above system for the following parameter values and function specifications:

$$f(x) = 68.4 \log(1 + X); \qquad \gamma = 0.1;$$

$$\delta = 0.8; \quad w = 35; \quad n = 10;$$

$$\theta_1 = \theta_2 = 1/3;$$

using a variety of replacement rates (values of ρ) ranging from 0.005 to 0.050. To check for asymptotic stability, we evaluated the eigenvalues of the Jacobian of the two-dimensional system at each equilibrium point, and verified that the eigenvalues have negative real parts. For each replacement rate, there exist (at least) two equilibria, one consisting largely of defectors, and the other consisting mainly of a mix of enforcers and cooperators. Table A1 provides the numerical details. Eigenvalues at the C–E type equilibria are real and negative, those at the D-equilibria are complex with negative real parts for sufficiently low replacement rates. In all cases, both types of equilibria are asymptotically stable.

The intuition for these results is the following. Near $\mathcal{S}$, the first term on the right-hand side of (A1) is negative and small in absolute value, since s_2 is small. So sufficiently close to $\mathcal{S}$, for positive ρ and θ_2, the selection pressure against defectors will be balanced by drift (the second term on the right-hand side of (A1)) in their favor. Now in this region, selection pressure will favor cooperators over enforcers. For the existence of a stable equilibrium, drift will have to favor enforcers over cooperators. This will happen if the frequency of enforcer births, $\theta_3 = 1 - \theta_1 - \theta_2$, is sufficiently large relative to θ_1, the frequency of cooperator births; that is, if $\theta_1 - s_1$ is less (more negative) than $\theta_3 - s_3$ by enough so as to offset the higher growth rate of s_1 coming from selection pressure. Since we have the further restriction that at the equilibrium, $s_1 < \bar{a}$, the end-point of $\mathcal{S}$, this means that when $\bar{a}$ is large, there is a wider range of values of θ_1 and θ_3 for which an equilibrium can exist. While this could be derived analytically, checking stability would still be possible only for numerical examples, since that involves signing the eigenvalues of the Jacobian of the system (A1).

If, instead of deterministic drift, we allowed for small stochastic perturbations along the lines of H. Peyton Young (1993) or Michihiro Kandori et al. (1993) there would be transitions from the neighborhood of one equilibrium to that of another. However, Glenn Ellison (1993) has shown that when the payoffs of the players depend on the strategies followed by the entire population (as is the case in this paper) rather than on those followed by a small subset of the population, then these transitions will be very infrequent. Given the timescales relevant for this paper, the introduction of stochastic perturbations is therefore unlikely to affect our main inferences.

REFERENCES

Acheson, James M. *The lobster gangs of Maine*. Hanover: University Press of New England, 1988.

———. "Capturing the Commons: Legal and Illegal Strategies," in Terry L. Anderson and Randy T. Simmons, eds., *The political economy of customs and culture: Informal solutions to the commons problem*. Lanham: Rowman and Littlefield, 1993.

Axelrod, Robert. "An Evolutionary Approach to Norms." *American Political Science Review*, December 1986, *80*(4), pp. 1095–111.

Banerjee, Abhijit V. and Weibull, Jörgen. "Evolutionary Selection and Rational Behavior," in Alan Kirman and Mark Salmon, eds., *Rationality and learning in economics*. Oxford: Basil Blackwell, 1994.

Bendor, Jonathan and Mookherjee, Dilip. "Institutional Structure and the Logic of Ongoing Collective Action." *American Political Science Review*, March 1987, *81*(1), pp. 129–54.

Benhabib, Jess and Radner, Roy. "The Joint Exploitation of a Productive Asset: A Game-Theoretic Approach." *Economic Theory*, April 1992, *2*(2), pp. 155–90.

Björnerstedt, Jonas and Weibull, Jörgen. "Nash Equilibrium and Evolution by Imitation." Discussion paper, Department of Economics, Stockholm University, 1994.

Bromley, Daniel W., ed. *Making the commons work*. San Francisco: ICS Press, 1992.

Chichilnisky, Graciela. "North South Trade and the Global Environment." *American Economic Review*, September 1994, *84*(4), pp. 851–74.

Ciriacy-Wantrup, Siegfried V. and Bishop, Richard C. " 'Common Property' as a Concept in Natural Resources Policy." *Natural Resources Journal*, 1975, *15*, pp. 713–27.

Clark, Colin W. *Mathematical bioeconomics: The optimal management of renewable resources*, 2nd. Ed. New York: John Wiley, 1990.

Cordell, John and McKean, Margaret A. "Sea Tenure in Bahia, Brazil," in Daniel W. Bromley, ed., *Making the commons work*. San Francisco: ICS Press, 1992.

Dasgupta, Partha. *An inquiry into well-being and destitution*. Oxford: Oxford University Press, 1993.

Dasgupta, Partha and Heal, Geoffrey. *Economic theory and exhaustible resources*. Cambridge: Cambridge University Press, 1979.

Dutta, Prajit K. and Sundaram, Rangarajan K. "The Tragedy of the Commons?" *Economic Theory*, July 1993, *3*(3), pp. 413–26.

Ellison, Glenn. "Learning, Local Interaction, and Coordination." *Econometrica*, September 1993, *61*(5), pp. 1047–71.

Fudenberg, Drew and Maskin, Eric. "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information." *Econometrica*, May 1986, *54*(3), pp. 533–54.

Gale, John; Binmore, Kenneth G. and Samuelson, Larry. "Learning to Be Imperfect: The Ultimatum Game." *Games and Economic Behavior*, January 1995, *8*(1), pp. 56–90.

Gordon, H. Scott. "The Economic Theory of a Common Property Resource: The Fishery." *Journal of Political Economy*, April 1954, *62*(2), pp. 124–42.

Guckenheimer, John and Holmes, Philip. *Nonlinear oscillations, dynamical systems, and bifurcations of vector fields*. New York: Springer Verlag, 1983.

Hardin, Garrett. "The Tragedy of the Commons." *Science*, December 1968, *162*, pp. 1243–48.

Hirshleifer, David and Rasmusen, Eric. "Cooperation in a Repeated Prisoners' Dilemma with Ostracism." *Journal of Economic Behavior and Organization*, August 1989, *12*(1), pp. 87–106.

Kandori, Michihiro; Mailath, George J. and Rob, Rafael. "Learning, Mutation, and Long Run Equilibria in Games." *Econometrica*, January 1993, *61*(1), pp. 29–56.

Martin, Fenton. *Common-pool resources and collective action: A bibliography*, Vols. 1 and 2, Workshop in Political Theory and Policy Analysis. Bloomington: Indiana University, 1989 and 1992.

Ostrom, Elinor. *Governing the commons: The evolution of institutions for collective action*. Cambridge: Cambridge University Press, 1990.

Ostrom, Elinor; Walker, James and Gardner, Roy. "Covenants With and Without a Sword: Self-Governance Is Possible." *American Political Science Review*, June 1992, *86*(2), pp. 404–17.

Putnam, Robert D. *Making democracy work*. Princeton, NJ: Princeton University Press, 1993.

Samuelson, Larry and Zhang, Jianbo. "Evolutionary Stability in Asymmetric Games." *Journal of Economic Theory*, August 1992, *57*(2), pp. 363–91.

Sethi, Rajiv. "Evolutionary Stability and Social Norms." *Journal of Economic Behavior and Organization*, January 1996, *29*(1), pp. 113–40.

Somanathan, E. "Deforestation, Property Rights and Incentives in Central Himalaya." *Economic and Political Weekly*, January 26, 1991, pp. PE37–46.

——. "Three Essays in the Economics of Institutions." Ph.D. dissertation, Harvard University, 1995.

Taylor, Lawrence J. "Community and Common Property in an Irish Fishing Settlement," in Bonnie J. McCay and James M. Acheson, eds., *The question of the commons*. Tucson: University of Arizona Press, 1987.

Taylor, Peter and Jonker, Leo. "Evolutionarily Stable Strategies and Game Dynamics." *Mathematical Biosciences*, July 1978, *40*(2), pp. 145–56.

van Damme, Eric. *Stability and perfection of Nash equilibria*. Berlin: Springer-Verlag, 1987.

Watts, Alison. "On the Uniqueness of Equilibrium in Cournot Oligopoly and Other Games." *Games and Economic Behavior*, forthcoming.

Weissing, Franz and Ostrom, Elinor. "Irrigation Institutions and the Games Irrigators Play: Rule Enforcement on Government- and Farmer-Managed Systems," in Fritz W. Scharpf, ed., *Games in hierarchies and networks: Analytical and empirical approaches to the study of governance institutions*. Frankfurt: Campus Verlag, 1993.

Weitzman, Martin L. "Free Access vs. Private Ownership as Alternative Systems for Managing Common Property." *Journal of Economic Theory*, June 1974, *8*(2), pp. 225–34.

Young, H. Peyton. "The Evolution of Conventions." *Econometrica*, January 1993, *61*(1), pp. 57–84.

Zeeman, Eric C. "Population Dynamics from Game Theory," in Zbigniew Nitecki and Clark Robinson, eds., *Global theory of dynamical systems*. Berlin: Springer Verlag, 1979.