

Cooperation in the Prisoner's Dilemma with Anonymous Random Matching



Glenn Ellison

The Review of Economic Studies, Vol. 61, No. 3. (Jul., 1994), pp. 567-588.

Stable URL:

<http://links.jstor.org/sici?sici=0034-6527%28199407%2961%3A3%3C567%3ACITPDW%3E2.0.CO%3B2-L>

The Review of Economic Studies is currently published by The Review of Economic Studies Ltd..

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/resl.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact support@jstor.org.

Cooperation in the Prisoner's Dilemma with Anonymous Random Matching

GLENN ELLISON
Harvard University

First version received October 1991; final version accepted February 1994 (Eds.)

The paper considers the repeated prisoner's dilemma in a large-population random-matching setting where players are unable to recognize their opponents. Despite the informational restrictions cooperation is still a sequential equilibrium supported by "contagious" punishments. The equilibrium does not require excessive patience, and contrary to previous thought, need not be extraordinarily fragile. It is robust to the introduction of small amounts of noise and remains nearly efficient. Extensions are discussed to models with heterogeneous rates of time preference and without public randomizations.

1. INTRODUCTION

Since the earliest work on the Folk Theorem, it has been well known that when two players face each other in a repeated prisoner's dilemma the "cooperative" outcome can be sustained as an equilibrium (Friedman (1971), Aumann and Shapley (1976)). A variety of extensions are possible. Given additional assumptions, the Folk Theorem has been shown to apply to N -player games, finite-horizon games of incomplete information, and games with imperfect observations (Fudenberg and Maskin (1986), Fudenberg, Levine, and Maskin (1993)).

In models of social games in which a large population of players are randomly matched it is reasonable to assume that players have limited information about other players' actions, e.g. players may observe only the outcome of matches in which they are personally involved. The results cited above are then not applicable. This paper follows those of Kandori (1992), Harrington (1991), and Okuno-Fujiwara and Postlewaite (1990) in investigating the extent to which Folk Theorem-type results may be obtained despite the special information structures generated by random matching games. In particular, I consider a random matching version of the prisoner's dilemma under the most extreme informational restriction—that players not only do not observe the outcomes of games in which they are not involved, but also are completely anonymous in that they can neither recognize nor communicate the identity of any of their past opponents. The main conclusion of this paper is that cooperation is possible in equilibrium and that this cooperation is somewhat robust. I hope that the argument is interesting to game theorists as an illustration of the variety of dynamics which may emerge in equilibrium as a response to informational limitations. The existence of a cooperative equilibrium also has practical implications in relation to several areas of recent research.

With anonymous random matching, it is impossible to maintain cooperation in a repeated game simply by punishing players who deviate. This paper builds on the observation of Kandori (1992) and Harrington (1991) that cooperation may nonetheless be possible if players use strategies with "contagious" punishments. In such strategies, when one player cheats in period t , his period- t opponent cheats from period $t+1$ on, infecting another player who cheats from period $t+2$ on, etc. For any fixed population size, Kandori provides an example of a game in which cooperative repeated-game equilibria exist, showing that we can define payoffs for the prisoner's dilemma which allow cooperation in a sequential equilibrium. However, when the population is large the argument applies only to games with extreme payoffs.

In this paper, I build on Kandori's arguments to study two main problems. First, for general payoffs in the prisoner's dilemma, is cooperation possible in a sequential equilibrium? I find that the answer is yes for sufficiently patient players. The argument demonstrates the versatility of contagious punishments which lead to a breakdown of cooperation after a single deviation. I assume at first that a publicly observable random variable is available. The public randomization allows the severity of the punishments to be easily adjusted so that the players fear a breakdown enough that they will not deviate first and destroy cooperation, but do not fear the breakdown so much that they are unwilling to contribute to its spread once it has begun. At several points I emphasize that this cooperation does not require unduly patient players.

The second problem is a study of the stability and efficiency of the equilibrium in a world with noise. Kandori observed that in the equilibrium he constructs a single deviation causes a permanent end to cooperation and comments that this fragility may make the equilibrium inappropriate as a model for trade. His observation reflects two quite distinct concerns. The first is a modeling issue I will refer to as stability. If we intend for the equilibrium to model cooperation in actual social settings and believe that in the real world punishments never last infinitely long we would like to construct an equilibrium with this property. Given public randomizations, this is not difficult. The second is a desire for a model which retains its efficiency in a world with noise. If we introduce noise by assuming that players either tremble and accidentally play the wrong strategy or misinterpret the actions of others, the equilibrium Kandori gives will be inefficient. Because cooperation eventually breaks down, the expected payoff to very patient players will be near the non-cooperative level. In the standard repeated prisoner's dilemma with noise, the results of Fudenberg, Levine and Maskin (1993) imply that this inefficiency can be avoided. In the random-matching model here I am able to show that for sufficiently small probabilities of mistakes being made there is a sequential equilibrium in which players need not change their strategies in response to the presence of mistakes, and in which the inefficiency is small. While this is clearly a limiting result, we can conclude that the cooperative equilibrium with anonymous matching need not be as fragile as has been portrayed.

While public randomizations are appropriate for many social situations, it is in the spirit of this paper to make do with as little information as possible. For this reason I also consider the problem of eliminating the reliance on public randomizations, finding that a cooperative equilibrium still exists. Interestingly, play in this equilibrium follows an unusual pattern with punishments scattered among intervening periods of cooperation. Payoffs in a model with noise are nearly efficient even though the equilibrium is no longer stable.

The questions analyzed here may be of interest in connection with several lines of research. First, in experimental economics it is a well-recognized concern that subjects who are asked to play a game several times may treat the situation as a repeated game.

To avoid repeated-game effects it is common practice to randomly match the players in an anonymous setting so that pairs of players do not meet repeatedly. The results here suggest, however, that given moderate population sizes random matching may not solve the problem.

Second, random-matching models have proven useful in recent studies of the economic institutions of trade. Greif (1989) discusses the Maghribi traders, a group of North African Jews who conducted trade in many Mediterranean countries in the 11th century. Milgrom *et al.* (1990) discuss trade in cities and fairs in Medieval Europe. In each case, the underlying model is one of a large number of traders who in each period are randomly paired with a trading partner. Each pair is presumed to play a game like the prisoner's dilemma with each party having both the opportunity and a private incentive to cheat the other by under-reporting sales on consignment, renegeing on promises to make future payments or deliveries, supplying goods of inferior quality, etc.

In this literature, institutions are seen as a way of avoiding the inefficiency of non-cooperative equilibria. Specifically, it is noted that the standard Folk Theorem equilibria of repeated games make informational demands which are unreasonable in a large society. Greif (1989) argues that the closeness of the Maghribi community did allow the necessary information exchange. He cites evidence that many traders maintained ties to traders in other cities. Via this network of relationships they would quickly learn the identity of any cheaters, allowing the offending parties to be punished. Milgrom *et al.* (1990) argue that such closeness no longer existed with the development of larger towns and trade fairs, and that this problem was resolved by the development of the Law Merchant, a private legal code whereby disputes could be tried before a judge who often lacked the power of enforcement. That I find cooperation to be possible in equilibrium without any institutions implies that it is more difficult to justify any institution as the least costly method of avoiding inefficiency. Note that it is certainly not claimed here that we would have expected to see cooperation without institutions, nor even that the equilibrium described here was feasible in Medieval trade fairs. I claim only that a consideration of whether it was feasible is necessary, and should sharpen our understanding of the role of the observed institutions.

Finally, several authors have explored the possibility that large population models may be used to reduce the multiplicity of equilibria in repeated games. Rosenthal (1979) discusses "rational Markovian hypotheses" in which all players react to steady-state conjectures based only on their current opponent's play in the previous period, not on any further history. In the case of the prisoner's dilemma, both players cheating in every period is the only such equilibrium (except in one special case). Green (1980) and Sabourian (1990) discuss models with noisy observation of an aggregate statistic and show that as the number of players grows large the equilibrium set shrinks to the static Nash outcome. With an information structure like that of this paper, Milgrom *et al.* (1990) note that with infinite population and an extreme matching rule where no player can affect his future opponents' play in any way, cheating is the only Nash equilibrium outcome. Our results suggest that large populations may do little to reduce the equilibrium set unless special assumptions are made.

The paper is organized as follows. Section 2 describes the model more precisely and exhibits a sequential equilibrium which sustains cooperation. Section 3 discusses the problem of stability and also shows that after introducing noise into the model we can still construct an equilibrium whose payoff approaches the efficient level as the amount of noise tends to zero, even for very patient players. Section 4 discusses the extension of the results to a model without public randomizations.

2. THE RANDOM-MATCHING MODEL

For the remainder of this paper, I analyze the model described below. The game has M players indexed by $i \in \{1, 2, 3, \dots, M\}$ where $M \geq 4$ is an even number. In each time period $t \in \{1, 2, 3, \dots\}$, the players are randomly matched into pairs with player i facing player $o_i(t)$. It is assumed that the pairings are independent over time and uniform so that

$$\text{Prob}\{o_i(t)=j|h_{t-1}\} = \frac{1}{M-1} \quad \forall j \neq i$$

for all possible histories h_{t-1} . At time t , each pair of players plays the prisoner's dilemma as shown below. The payoff g is taken to be positive with l non-negative so that each player has D as a dominant strategy in the stage game. All players have discount factor $\delta \in (0, 1)$ and their payoffs are the discounted sum of the payoffs in each stage game. At the end of period t , each player observes only the outcome of the prisoner's dilemma he and his opponent played. He does not observe the identity $o_i(t)$ of his opponent and does not observe the outcome of any of the games played by other pairs of players.

	C	D
C	1, 1	$-l, 1+g$
D	$1+g, -l$	0, 0

In addition, I will assume in this section and in the one which follows that before players choose their actions in period t , they observe a public random variable q_t which is drawn independently from a uniform distribution on $[0, 1]$. In some situations, it seems reasonable to assume that such a randomization is available. For example, all traders at a market may have access to the same newspaper or hear the same government announcements. In any case, the use of public randomizations simplifies the exposition below. I will later discuss how many of the same results can be obtained without public randomizations.

The first thing to note about this model is that we can not implement the types of strategies usually used to prove the Folk Theorem. For example, when a player is the first to deviate, there is no way of identifying him, so it will be impossible to punish one player more severely and reward others for carrying out the punishment. Also, there is no obvious way to convey any information about the precise time of the deviation so that players could coordinate on something like T -period punishments.

Kandori (1992) shows that contagious punishments can be used to sustain collusion in some circumstances. Specifically, he shows that for any population size M , we can choose the payoff l so that cooperation is a sequential equilibrium for sufficiently patient players. The choice of l is used to give players an incentive to carry out the punishment which follows a deviation. Unfortunately, the value of l Kandori uses grows without bound as M increases and may be unreasonable for moderate values of M .

The main result of this section is that cooperation is indeed a sequential equilibrium of the random-matching game for any payoffs g and l . The equilibrium is supported by strategies like Kandori's which rely on contagious punishments. All subsequent results will rely on similar strategies. The following proposition gives the basic result.

Proposition 1. *Consider the random-matching model with public randomizations described above where $M \geq 4$ players play the prisoner's dilemma with $g > 0, l \geq 0$. Then,*

$\exists \delta < 1$ such that $\forall \delta \in [\delta, 1)$ there is a sequential equilibrium $s^*(\delta)$ of the repeated game in which all players play C in every period along the equilibrium path.

Before giving a formal proof, let me first discuss the strategies $s^*(\delta)$ which will support the equilibrium. The strategies described below employ a contagious process by which a deviation in period t will usually lead to two players playing D in period $t+1$, then four players playing D in period $t+2$, etc. The result is a breakdown of social cooperation which punishes all players after one deviates. Given a function $q(\delta)$ to be defined below, the strategies are as follows.

In period 1, all players begin play according to phase I.

Phase I. Play C in period t .

If (C, C) is the outcome for matched players i and j , both play according to phase I in period $t+1$.

If (C, D) , (D, C) , or (D, D) results in the game between players i and j , then at time $t+1$ both play according to phase II if $q_{t+1} \leq q(\delta)$ and according to phase I if $q_{t+1} > q(\delta)$.

Phase II. Play D in period t .

In period $t+1$ play according to phase I if $q_{t+1} > q(\delta)$ and according to phase II if $q_{t+1} \leq q(\delta)$.

The public randomizations are being used to adjust the severity of the punishment phase so that it lasts $1/(1-q(\delta))$ periods on average. The basic idea of the proof is this. In a sequential equilibrium the continuation payoffs of the players must satisfy two constraints derived from players not having a profitable single-period deviation. First, players must not want to deviate and play D in phase I. When punishments are of infinite duration (i.e. for $q(\delta) = 1$), sufficiently patient players will not want to cause a breakdown of cooperation in phase I so this constraint is satisfied. Second, we must recognize that in phase II players might deviate and play C in hopes of slowing the spread of the contagious punishment. When punishments almost never continue (i.e. for $q(\delta) \approx 0$) there is no possible gain to deviating in phase II so this constraint will be satisfied.

To prove the proposition, I show that there exists at least one value $q(\delta)$ which is both large enough to prevent deviations in phase I and small enough to prevent deviations in phase II. The intuitive reason why this can be done is simple. In either phase I or phase II player i gets the same short-term gain of g from playing D when this opponent cooperates. However, starting a punishment by playing D in phase I causes a greater loss in continuation payoff than does spreading a punishment by playing D in phase II. Once play is in phase II, cooperation is breaking down anyway so one extra deviation has limited impact. Choosing an appropriate punishment severity, the loss from starting a punishment deters playing D in phase I, but the loss from spreading a punishment does not deter playing D in phase II.

To formalize this argument let k be the number of players who are playing according to phase II at the start of period t . Let $f(k, \delta, q)$ be player i 's (per period) continuation payoff from period t on when all players are playing the strategies above, and player i and $k-1$ others are playing according to phase II. If player i deviates and plays D in phase I in period t , he gains g in period t but will have a lower continuation payoff from

period $t+1$ on. To show that no deviation is profitable in phase I we must show that

$$(1-\delta)g \leq \delta q(\delta)(1-f(2, \delta, q(\delta))). \quad (1)$$

We can also derive a similar sufficient condition for there to be no profitable deviation in phase II. If player i deviates and plays C in phase II at time t we have one of two possibilities. First, he could be matched with someone else who is playing according to phase II. In this case, the result in period t is (C, D) instead of (D, D) , and continuation payoffs are unaffected. Clearly, player i is not better off because $l \geq 0$. Second, player i might be matched with someone who is playing according to phase I. The period t outcome is then (C, C) instead of (D, C) so player i loses g in period t . In the continuation game, however, one fewer player will be playing according to phase II. The deviation is not profitable if

$$(1-\delta)g \geq \delta q(\delta)E_j[f(j, \delta, q(\delta)) - f(j+1, \delta, q(\delta))]$$

where the expectation reflects player i 's beliefs over the number of players who will play according to phase II at time $t+1$. To show that this relation holds for beliefs of player i consistent with sequential equilibrium, a sufficient condition is to show that it holds pointwise, i.e.

$$(1-\delta)g \geq \delta q(\delta)(f(j, \delta, q(\delta)) - f(j+1, \delta, q(\delta))) \quad \forall j \geq 3. \quad (2)$$

(The beliefs must assign probability zero to $j \leq 2$ because when player i is in phase II, his opponent when he first saw cheating and that player's period t opponent will also play according to phase II in period $t+1$.) When (1) and (2) hold, we have a sequential equilibrium. In establishing these relations, both the result and the intermediate calculations of the following lemma will prove useful.

Lemma 1. $f(k, \delta, q)$ is convex in k for $k \geq 1$, i.e.

$$f(k, \delta, q) - f(k+1, \delta, q) \geq f(k+s, \delta, q) - f(k+s+1, \delta, q)$$

for all $s \geq 1$.

The lemma simply states that the loss in continuation payoff from having one extra player infected declines as the number of infected players grows. This is to be expected as, when many players are infected, the one extra player not infected in period t is likely to become infected in period $t+1$ anyway and thus never have a chance to affect player i 's payoff. The proof is straightforward once I introduce enough notation.

Proof of Lemma 1. Note that

$$f(k, \delta, q) = E_{\omega}g(k, \delta, q, \omega),$$

where ω is the random variable whose realization is a pairing of all players in each period, and the function g gives player 1's continuation payoff for a given matching when he and players $2, \dots, k$ are playing according to phase II. For expositional convenience I define $h(k, \delta, q, \omega)$ to be player i 's continuation payoff when he and players $2, \dots, k$ and player M are playing according to phase II. Clearly

$$E_{\omega}g(k+1, \delta, q, \omega) = E_{\omega}h(k, \delta, q, \omega).$$

I show that

$$E_\omega[g(k, \delta, q, \omega) - h(k, \delta, q, \omega)] \geq E_\omega[g(k+s, \delta, q, \omega) - h(k+s, \delta, q, \omega)]$$

by showing that the inequality holds for every realization of ω .

Define the set $C(t, k, \omega)$ by

$$C(0, k, \omega) = \{k+1, k+2, \dots, M\},$$

$$C(t+1, k, \omega) = \{i \in C(t, k, \omega) \mid o_i(t, \omega) \in C(t, k, \omega)\}.$$

$C(t, k, \omega)$ will be the set of players still playing according to phase I in period t when $q_s \leq q$ for all $s \leq t$ and players $1, 2, \dots, k$ begin in phase II in period 0.

Define the set $D(t, \omega)$ by

$$D(0, \omega) = \{M\}$$

$$D(t+1, \omega) = D(t, \omega) \cup \{i \mid o_i(t, \omega) \in D(t, \omega)\}.$$

$D(t, \omega)$ gives the set of all players who will be playing according to Phase II in period t when $q_s \leq q$ for all $s \leq t$ and player M begins in phase II in period 0.

Note that the payoff to player 1 in period t differs between the situations of $g(k, \delta, q, \omega)$ and $h(k, \delta, q, \omega)$ only if $q_s \leq q$ for all $s \leq t$ and only if his opponent $o_1(t, \omega)$ plays C when players $1, 2, \dots, k$ start in phase II but plays D when players $1, 2, \dots, k$, and player M start in phase II. Thus,

$$g(k, \delta, q, \omega) - h(k, \delta, q, \omega) = \sum_{t=0}^{\infty} (1-\delta)q^t \delta^t (1+g)I(o_1(t, \omega) \in C(t, k, \omega) \cap D(t, \omega)). \quad (3)$$

(The notation $I(E)$ indicates a function equal to one or zero depending on whether the deterministic condition E is true or false.) The definition of C clearly implies that

$$C(t, k+s, \omega) \subset C(t, k, \omega)$$

so

$$C(t, k+s, \omega) \cap D(t, \omega) \subset C(t, k, \omega) \cap D(t, \omega)$$

and the expansion (3) gives the desired result. \parallel

We are now in a position to give

Proof of Proposition 1. Let $s^*(\delta)$ be the strategy profile given above. It suffices to demonstrate the existence of a $\underline{\delta} < 1$ such that (1) and (2) hold for all $\delta \in [\underline{\delta}, 1)$. To establish the relation (1), we will simply define $\underline{\delta}$ and $q(\delta)$ on $[\underline{\delta}, 1)$ so that (1) holds with equality. To see that this is possible, note that for $q(\delta) = 1$, punishments are infinite so

$$f(2, \delta, 1) = (1-\delta) \sum_{t=0}^{\infty} \delta^t a_t$$

where a_t is the expected payoff in the t -th period after phase II play begins. With probability 1 all players will eventually be infected and start playing D so $a_t \rightarrow 0$. We then have

$$\lim_{\delta \rightarrow 1} \frac{\delta}{1-\delta} (1-f(2, \delta, 1)) = \infty,$$

$$\lim_{\delta \rightarrow 0} \frac{\delta}{1-\delta} (1-f(2, \delta, 1)) = 0.$$

By continuity we can choose $\underline{\delta} \in (0, 1)$ so that

$$\frac{\underline{\delta}}{1-\underline{\delta}} (1-f(2, \underline{\delta}, 1)) = g. \quad (4)$$

Note that when (1) holds with equality, a player in phase I is exactly indifferent between playing C and D . The payoff to a player who plays D in period 1 is $f(1, \delta, q(\delta))$. Thus, (1) holds with equality only if

$$\frac{\delta q(\delta)}{1-\delta} (f(1, \delta, q(\delta)) - f(2, \delta, q(\delta))) = g. \quad (5)$$

The converse is also true. When (5) holds, a player in phase I is exactly indifferent between playing D in period t (and following the equilibrium strategies thereafter) and playing C in period t then deviating and playing D in period $t+1$. Applying the same indifference again, he is also indifferent between deviating in period t and playing C in periods t and $t+1$ and then deviating in period $t+2$. Repeating this process, he is indifferent between deviating in period t and cooperating in all future periods. This implies that (1) holds with equality.

From expansion (3) we have

$$\begin{aligned} & \frac{\delta q}{1-\delta} (f(k, \delta, q) - f(k+1, \delta, q)) \\ &= \sum_{t=0}^{\infty} (\delta q)^{t+1} (1+g) \text{Prob} \{o_1(t, \omega) \in C(t, k, \omega) \cap D(t, \omega)\}. \end{aligned} \quad (6)$$

As the right-hand side depends only on the product δq , we simply define

$$q(\delta) \equiv \underline{\delta} / \delta.$$

Then, for any $\delta \in [\underline{\delta}, 1)$ $\delta q(\delta) = \underline{\delta}$ and

$$\frac{\delta q(\delta)}{1-\delta} (f(1, \delta, q(\delta)) - f(2, \delta, q(\delta))) = g$$

as desired.

The result of the lemma now immediately implies (2) and hence completes the proof. \parallel

At this point, we may better understand the role of the prisoner's dilemma in the argument above. In the prisoner's dilemma, the maximum one-period gain from cheating is identical to the short-term loss a player incurs by not playing the static Nash equilibrium when he successfully slows a punishment. The convexity argument above establishes that the loss starting a punishment is greater than the gain from slowing a punishment, and therefore that the short-term loss/gain from following the equilibrium strategies in phase I/II can be made to lie between these future effects. In general games, such an argument shows that the symmetric strategy profile A is an equilibrium if the payoff $u(A, A)$ dominates the payoff $u(s^*, s^*)$ of a Nash equilibrium and s^* is a best response to A . (This, for example, gives a Folk Theorem for games with a dominant-strategy equilibrium.) In other games there may be a much more profitable deviation from cooperation, and hence strategies similar to those described above will sustain a Nash equilibrium only if there is a sufficient difference between the loss from starting a punishment and the gain from

TABLE I

Discount factor sufficient to maintain cooperation

	$g=0.01$	$g=1$	$g=10$
$M=2$	0.01	0.50	0.91
$M=4$	0.03	0.68	0.95
$M=10$	0.08	0.79	0.97
$M=100$	0.35	0.89	0.985
$M=1000$	0.57	0.93	0.990

slowing a punishment. To establish a more general Folk Theorem, we would need somehow to create punishments for which that difference could be made arbitrarily large.

It should also be noted that the assumption that the random matching is uniform has been made largely for convenience. With this assumption the players have symmetric continuation payoffs and the analysis is simplified by the fact that these can be written as a function of the number of players in each phase. The idea that the long-term consequences of a single deviation are smaller when some players are already in the punishment phase appears to be much more general, and might be applicable to populations with other matching rules such as the local matching rules discussed in Ellison (1993). The argument does rely, however, on the matching being sufficiently symmetric so that all players have the appropriate incentive to avoid or to spread punishments.

In a full-information model, the "grim" strategies immediately punish a player who has cheated once. In contrast, the contagious punishment takes time to spread throughout the population so that a player may be able to cheat several opponents before he begins to suffer from the punishment phase he has brought on. This observation leads us to ask whether the equilibrium described in Proposition 1 requires undue patience on the part of the players.

Table I gives the minimum value of δ which can sustain cooperation for several population sizes M and for several values of the gain g to deviation. For comparison, I have also listed under the heading $M=2$ the discount factor necessary for the standard "grim" equilibrium in a two-player game. In a limiting sense, the behaviour of our model matches that of Green and Sabourian. For a fixed discount factor, cooperation will be impossible if the population size is sufficiently large. From the table, however, we can see that for reasonable population sizes patience is simply not a problem for our equilibrium. With the extreme gain ($g=10$), cooperation is possible in a population of one thousand players if players meet one opponent per month and discount the future at a rate of 5% per year. For the more standard payoffs with $g=1$, cooperation is possible in the same population even if players meet only one opponent per year. Moreover, a more detailed look at the numbers in the table suggests that (as would be expected given the exponential growth of contagious punishments) doubling the frequency with which players meet squares the size of the population for which a cooperative equilibrium exists. Although I have not done the calculation, this would imply that cooperation is possible in any population of fewer than a trillion players (with $g=1$) if each player meets one opponent per quarter.

To better appreciate the power of the contagious punishments in large populations, it is instructive to compare the discount factors of Table I to those necessary for another large population equilibrium. While this paper focusses on completely anonymous matching, for some applications it may be reasonable to make the less stringent assumption that identities can be observed but not communicated. In such a model, we could sustain cooperation via personal retaliation strategies where a deviation by player i in period t

causes his period t opponent $o_i(t)$ to play D whenever they are matched in the future. Note that this equilibrium requires frequent *individual* interactions, and thus requires far more patient players than does the equilibrium with contagious punishments. If player i cheats in period t , he gains g in period t and loses 1 in each future period in which he is again matched with $o_i(t)$. This gives a cooperative equilibrium only if

$$\sum_{t=1}^{\infty} \delta^t \frac{1}{M-1} \geq g \Leftrightarrow \delta \geq \frac{g(M-1)}{1+g(M-1)}.$$

For $g=1$ and $M=1000$, for example, this requires $\delta=0.999$, whereas $\delta=0.93$ is sufficient for the equilibrium with contagious punishments.

3. STABILITY AND EFFICIENCY WITH NOISE

The cooperative equilibrium described in Section 2 exhibits the desirable property of global stability described by Kandori (1992). That is, after *any* finite history, the continuation payoffs of the players eventually return to the cooperative level (with probability 1). Obviously, this is a result of the introduction of public randomizations. The stability does suggest, though, that robustness in this sense is not a big problem for this model.

A more interesting question is whether we can still sustain a nearly efficient outcome in a model with noise. Suppose we really believed that the model of Section 2 with its completely rational players and perfect observations were an accurate depiction of reality. Even if players follow the strategies of an equilibrium with infinite punishments, in equilibrium the punishment never begins, so we have no reason to care about the behaviour of the continuation payoffs after a deviation. On the other hand, suppose that there is noise in the model, as players either act irrationally some fraction of the time, or try to cooperate but make mistakes and play the wrong strategy or misinterpret their opponent's action. Again, I would argue that whether an equilibrium is stable is not the appropriate question to ask. If we have a globally stable equilibrium in which the continuation payoffs return to the cooperative level so slowly so that with noise the equilibrium has an expected payoff near zero, stability is not comforting. Suppose we have two different equilibria which have the same loss of efficiency after any deviation. Should we care if one equilibrium has all the inefficiency right away and then returns to cooperation while the other spreads out the same inefficiency over an infinite time period? The answer, I think, is that all that matters is the degree of efficiency the equilibrium attains in a model with the noise explicitly modelled.

In the two-player repeated prisoner's dilemma complete efficiency can be attained in the limit $\delta \rightarrow 1$ (Fudenberg, Levine, and Maskin (1993)). I will now introduce noise into the model of Section 2 by assuming that all players are constrained to play D with probability at least $\varepsilon > 0$ at every possible history. In the trade example, this could correspond to players trying to supply a high-quality good but accidentally supplying one which proves defective. A similar result could be obtained if we assumed instead that there was only noise in observing opponents' actions. While the equilibrium of Section 2 is not robust to this noise (because of the exact indifference during phase I play), the proposition below shows that for a slightly longer punishment length we do in fact have an equilibrium robust to this noise. While the existence of a fully efficient equilibrium is still an open question, the equilibrium described is approximately efficient in the sense that it approaches efficiency as $\varepsilon \rightarrow 0$.

Proposition 2. *Under the assumptions of Proposition 1, there exists $\delta' < 1$ and a set of strategy profiles $s^*(\delta)$ for $\delta \in [\delta', 1)$ of the random-matching game with the following three properties:*

1. *In the game with discount factor δ , $s^*(\delta)$ is a sequential equilibrium with all players playing C on the path in every period.*
2. *Define $s^*(\delta, \varepsilon)$ to be the strategy which at each history assigns probability ε to D and probability $1 - \varepsilon$ to the action given by $s^*(\delta)$. Then, there exists $\bar{\varepsilon} > 0$ such that $\forall \varepsilon < \bar{\varepsilon}$ $s^*(\delta, \varepsilon)$ is a sequential equilibrium of a perturbed game where all players are required to play D with probability at least ε at each history.*
3. *For u_i defined to player i 's expected per period payoff,*

$$\lim_{\varepsilon \rightarrow 0} \lim_{\delta \rightarrow 1} u_i(s^*(\delta, \varepsilon)) = 1.$$

Outline of Proof. We will show that $s^*(\delta)$ can be taken to have the same form as the strategy profile in the proof of Proposition 1, but with a slightly larger probability $q'(\delta)$ of continuing in a punishment phase. The proof requires attention to some tedious details, so I only outline the proof here and leave the rest to the Appendix.

To begin, I give a slight extension of Lemma 1, showing that the continuation payoff function f is strictly convex. The strict convexity allows us to choose a slightly larger $q'(\delta)$ so that the two inequalities which describe a player's loss from deviating in phase I or phase II of the model with no noise hold strictly. Formally, the Appendix shows that we can choose δ' , $q'(\delta)$ and $\eta > 0$ for which $\delta \geq \delta'$ implies

$$\frac{\delta q'(\delta)}{1 - \delta} (f(0, \delta, q'(\delta)) - f(2, \delta, q'(\delta))) > g + \eta \tag{7}$$

and

$$\frac{\delta q'(\delta)}{1 - \delta} (f(k, \delta, q'(\delta)) - f(k + 1, \delta, q'(\delta))) < g - \eta \quad \forall k \geq 2. \tag{8}$$

This immediately gives property 1.

To show that these strategies give an equilibrium for all sufficiently small ε requires two further steps. First, it must be shown that the left-hand side of each equation is uniformly continuous in ε so that for small enough ε the inequalities above still hold but with η replaced by $\eta/2$. For $f(k, \delta, q, \varepsilon)$ defined to be the continuation payoff of the strategies $s^*(\delta, \varepsilon)$ the Appendix demonstrates the existence of an $\bar{\varepsilon}_1 < 0$ such that for any $\varepsilon < \bar{\varepsilon}_1$

$$\frac{\delta q'(\delta)}{1 - \delta} (f(1, \delta, q'(\delta), \varepsilon) - f(2, \delta, q'(\delta), \varepsilon)) > g + \eta/2 \tag{9}$$

and

$$\frac{\delta q'(\delta)}{1 - \delta} (f(k, \delta, q'(\delta), \varepsilon) - f(k + 1, \delta, q'(\delta), \varepsilon)) < g - \eta/2 \quad \forall k \geq 2. \tag{10}$$

Second, we have a new complication in that when a player is playing according to phase I, he can no longer believe with probability 1 that all other players are doing so. Again though, as $\varepsilon \rightarrow 0$, this uncertainty also has an effect which vanishes so that the incentives to cooperate are maintained for sufficiently small ε . This completes the proof of 2.

Finally, the proof that we get efficiency in the limit is easy. The basic idea is that the punishment phases have a finite expected length bounded above by a constant independent of δ for δ close to 1. As $\varepsilon \rightarrow 0$ a vanishing fraction of the periods is spent in a punishment phase, so the expected payoff tends to the efficient level. Again the details are in the Appendix. ||

The results of Proposition 2 indicate that the equilibrium I have described is far less fragile than it might appear at first. The same strategies yield an equilibrium for all sufficiently small amounts of noise, so players can cooperate even if they do not know the precise frequency with which other players make mistakes. Further, the strategies are truly supporting cooperation in the sense of having nearly efficient payoffs with noise.

The theoretical notion of stability established in Proposition 2 has practical significance, for example, in that it implies that a formal justification of institutions cannot rely on the simplest limiting notion of robustness. I should note, however, that in other ways contagious equilibria may be very non-robust. For example, if we wish to assess whether a cooperative equilibrium was possible in a particular population of traders, we would want to consider reasonable mistake probabilities. Population size then becomes a major concern, because the $\bar{\varepsilon}$ defined in Proposition 2 may be extremely small and is decreasing in M . If even one player is expected to tremble in each period then mistakes will be far too common for players to want to cooperate in order to avoid starting a punishment. While robustness to large trembles is undoubtedly a problem, it is interesting to note that there are ways to modify the model to accommodate more frequent mistakes. First, we might suppose as in Ellison (1993) that we have some type of local matching rule in which each player is likely to meet only say 50 opponents regardless of the population size. Because a player suffers soon after starting a punishment in his neighbourhood and cares mostly about whether a punishment phase is ongoing nearby, it might be possible to construct an equilibrium in which $\bar{\delta}$ and $\bar{\varepsilon}$ can be chosen independently of the population size and $\bar{\varepsilon}$ is not so extremely small. Second, even with uniform matching we might (following Milgrom *et al.* (1990)) modify the stage game so that a player who accidentally trembles has the opportunity to give back his excess payoff at a trial and avoid the start of a punishment. If mistakes result from independent trembles at each information set, accidental punishments would now be much less likely. While a bit far-fetched, this does provide an alternative justification for legal institutions.

What is probably more important practically and harder to overcome is that the argument above deals only with trembles. If one player were "crazy" and always played D (or simply was unaware which equilibrium was being played) again the contagious strategies would not support cooperation. In large populations, the assumption that all players are rational and know their opponents' strategies may be both very important to the conclusions and fairly implausible.

Returning to our standard model, the fact that each action in our equilibrium with contagious punishments is a strict best response also allows the further extension that follows. In a large population, we may want to allow for heterogeneity among the players. In particular, it is probably reasonable to assume that the players have different rates of time preference. In each of the first two propositions, the equilibrium strategy profile $s^*(\delta)$ is a function of the discount factor. For each discount factor δ , the equilibrium involves a different probability $q(\delta)$ of continuing within the punishment phase. Hence, the strategies are only appropriate for a population of players all of whom share a common discount factor. As long as all of the players are sufficiently patient, however, we can eliminate this restriction. The proposition below guarantees the existence of a sequential equilibrium

strategy profile s^* which is not a function of δ . This profile will then sustain cooperation regardless of whether the population shares a common discount factor. For convenience, I shall discuss only a model without noise although the arguments clearly extend to the results of Proposition 2 as well. The proof is similar to that of Proposition 2, but is less involved.

Proposition 3. *Under the assumptions of Proposition 1, there exists a strategy profile s^* and a constant $\underline{\delta}'' < 1$ such that $\forall \delta \in [\underline{\delta}'', 1)$, s^* is a sequential equilibrium of the repeated matching game and all players play C in every period on the path of s^* .*

Proof. Once again, let s^* be a strategy profile like the one described in the proof of Proposition 1, but this time with punishment probability $q'' \equiv \lim_{\delta \rightarrow 1} q'(\delta)$. (The function $q'(\delta)$ is defined in the proof of Proposition 2. Note that q'' a constant independent of δ and that $q'' = \underline{\delta}'$, with $\underline{\delta}'$ the value chosen in that proof.) Intuitively, s^* is an equilibrium for δ close to one because when δ approaches one, q' approaches $q'(\delta)$ and the strategy profile s^* approaches the equilibrium $s^*(\delta)$ of Proposition 2. A formal proof along these lines is rather tedious and involves several limiting arguments like those given in the proof of Proposition 2. Instead, a simple constructive proof is given below.

Let $\underline{\delta}'' = \underline{\delta}/q''$, where $\underline{\delta}$ is as defined in (4). From (6) we know that $(\delta q / (1 - \delta))(f(1, \delta, q) - f(2, \delta, q))$ depends only on the product δq and is increasing in that expression. From $\delta \geq \underline{\delta}''$ we have $\delta q'' \geq \underline{\delta}$ and hence

$$\begin{aligned} \frac{\delta q''}{1 - \delta} (f(1, \delta, q'') - f(2, \delta, q'')) &\geq \frac{\underline{\delta}}{1 - \underline{\delta}} (f(1, \underline{\delta}, 1) - f(2, \underline{\delta}, 1)) \\ &= g, \end{aligned}$$

with the final equality following from (4). As in the proof of Proposition 1, this implies that no player expects to gain from a single-period deviation in phase I.

For any $\delta \in [\underline{\delta}'', 1)$, we also have $\delta q'' < q'' = \underline{\delta}'$. Hence, from (6), (8) and $q'(\underline{\delta}') = 1$ we know that

$$\begin{aligned} \frac{\delta q''}{1 - \delta} (f(2, \delta, q'') - f(3, \delta, q'')) &< \frac{\underline{\delta}'}{1 - \underline{\delta}'} (f(2, \underline{\delta}', 1) - f(3, \underline{\delta}', 1)) \\ &< g. \end{aligned}$$

This equation, combined with the convexity of f in its first argument implies that no player expects to gain from a single-period deviation from phase II play. \parallel

A potentially disturbing aspect of the preceding proof is that because it involves another limit as $\delta \rightarrow 1$, the equilibrium with heterogeneous discount factors might require far more patient players than was previously necessary. From Table I we know that Propositions 1 and 2 do not require unreasonably patient players. Certainly, the equilibrium described in Proposition 3 will sometimes require more patient players. This is particularly true when the gain g from deviation is small so that it is hard to get players to carry out punishments. For example, for a population of 100 players, if we take g to be 0.01 the equilibrium as constructed requires $\delta = 0.96$. Usually, though, we will think of g as being much larger. In the trade example, the payoff of 1 is the profit or consumer surplus from an honest transaction and g represent the additional profit from cheating (e.g. non-payment), which is liable to be at least as large as the profits from honest trade. For larger values of g , a cooperative equilibrium with heterogeneity often requires no

greater patience than was necessary for cooperation in the homogeneous population model. When the constraint that players be willing to carry out punishment is sufficiently far from binding, we can simply use infinite punishments for all $\delta \in [\underline{\delta}, 1)$ to get an equilibrium. Numerical calculations show this to be the case for each of the population sizes given in Table I for $g=1$ or $g=10$.

4. COOPERATION WITHOUT PUBLIC RANDOMIZATIONS

Throughout this paper, I have assumed that a public randomizing device is available. For many applications, including trade at a market, the assumption seems reasonable. Whenever all the players are present at the same physical location it seems likely that if the players looked hard enough they could find some random factor like the weather which everyone could observe and hence use to coordinate. Nonetheless, the focus of this paper is to describe how cooperation can be maintained with very little information available to the players. In this spirit then, I will discuss what can be done without public randomizations.

In Fudenberg and Maskin's (1986) proof of the perfect Folk Theorem, public randomizations play a crucial role in allowing the adjustment of players continuation payoffs necessary for maintaining exact indifference. Fudenberg and Maskin (1991) show that public randomizations, are, in fact, not necessary for this purpose. The crucial insight is that payoffs in the convex hull of the set of feasible payoffs can be obtained instead from a deterministic sequence of play.

In this paper, randomizations are playing two quite distinct roles. First, they are used as a coordinating device so that all players can simultaneously return to cooperation at the end of a punishment phase. The simultaneity is important because all players only slightly prefer cooperating when all others are doing so. If the probability that everyone else returns to cooperation in period t is not very close to one, no one will be willing to try returning to cooperative play. Coordination then allows the construction of a globally stable equilibrium. Whether global stability is possible without the public randomizations is unknown.

The second role of the public randomizations in this paper is to adjust the expected duration, and hence the severity of the punishment. This is the property which enabled us to construct strategies where punishments deter cheating, but are not so severe that individuals would be unwilling to carry them out. In the argument below, I show that for large enough discount factors it is possible to adjust the severity of the punishments in a completely different way—spreading out the punishments over time. This will allow us to establish the most important results of the paper even without the availability of public randomizations.

The ability to soften punishments by delaying them is at the heart of the following lemma. The lemma guarantees that any game which has a cooperative equilibrium for some interval of discount factors has a cooperative equilibrium for all discount factors near one as well. I hope that the very simple proof makes the lemma interesting in its own right.

Lemma 2. *Let $G(\delta)$ be any repeated game of complete information, and suppose that there is a non-empty interval (δ_0, δ_1) such that $G(\delta)$ has a sequential equilibrium $s^*(\delta)$ with outcome a for all $\delta \in (\delta_0, \delta_1)$. Then, there exists $\underline{\delta} < 1$ such that $\forall \delta \in (\underline{\delta}, 1)$ we can also define a strategy profile $s^{**}(\delta)$ which is a sequential equilibrium of $G(\delta)$ with outcome a .*

Proof. The key observations here is that for δ close enough to 1, we can simulate the situation of smaller discount factors by using slower responses.

Take $\bar{\delta} = \delta_0/\delta_1$. For any $\delta \in (\bar{\delta}, 1)$ there exists an integer $N(\delta)$ for which

$$\delta^{N(\delta)} \in (\delta_0, \delta_1).$$

When there is more than one such integer take $N(\delta)$ to be as large as possible. Now, have the players treat the game $G(\delta)$ as if it were $N(\delta)$ separate games, the first taking place in periods

$$1, N(\delta) + 1, 2N(\delta) + 1, 3N(\delta) + 1, \dots,$$

the second in periods

$$2, N(\delta) + 2, 2N(\delta) + 2, 3N(\delta) + 2, \dots,$$

etc. Just as is the case in finding Markov equilibria, if for some set T all other players play strategies in period t which do not depend on the outcomes in all periods $t' \in T$, then the best response for player i can be taken to be independent of the outcomes in all periods $t' \in T$ as well. Hence, to show that we have an equilibrium $s^{**}(\delta)$ for $G(\delta)$ it suffices to show:

1. The strategies $s^{**}(\delta)$ give play in period $aN + b$ which does not depend on play in period $cN + d$ if $(b - d)$ is not a multiple of N .
2. Restricting consideration to each "component" game played in periods

$$b, N(\delta) + b, 2N(\delta) + b, 3N(\delta) + b, \dots,$$

the restriction of the strategy profile $s^{**}(\delta)$ is a sequential equilibrium.

The obvious choice of $s^{**}(\delta)$ is to play the equilibrium $s^*(\delta^{N(\delta)})$ in each of the $N(\delta)$ component games described above. In our prisoner's dilemma example, this would mean that if player i or his opponent plays D in period $aN(\delta) + b$, player i plays D in periods

$$(a + 1)N(\delta) + b, (a + 2)N(\delta) + b, \dots,$$

but does not change his planned play in any other period. Within these component games, players have discount factor $\delta^{N(\delta)}$, so $s^*(\delta^{N(\delta)})$ satisfies the second condition. Clearly, we have a sequential equilibrium. ||

Note that when $q'(\delta) = 1$, the strategies described in the proof of Proposition 2 prescribe infinite punishments, and hence do not require randomizations. In particular, $\bar{\delta}'$ was defined so that taking $q = 1$ gives a sequential equilibrium. In order to apply Lemma 2, we need only show that infinite punishments also yield a sequential equilibrium for a small interval of discount factors around $\bar{\delta}'$. This result is not hard. It is simply another application of the fact that each action is a strict best response. The resulting equilibrium of the game has a peculiar appearance with punishments being softened by being delayed into the future, spread among intervening periods of cooperation. In the trade example, this might mean that if a single deviation occurs on a Friday, eventually we will see all players cheating on every third Friday but cooperating on all other days. The punishments are of infinite duration so with noise, eventually all players will cheat in all periods. Despite this, the punishments are still no more severe than the punishments of the previous section. As players become more patient, the punishment periods become correspondingly further apart. The somewhat surprising result is that in the limit as the amount of noise vanishes, the equilibrium approaches efficiency. These results are summarized below.

Proposition 4. *The results of Proposition 2 still hold in a model where no public randomizations are available.*

Proof. In order to establish the first two results of Proposition 2, that there is a sequential equilibrium which remains an equilibrium for sufficiently small amounts ε of noise, it will suffice to show that for a fixed range of discount factors the standard strategies with $q=1$ give a sequential equilibrium. Just as in Proposition 3, we apply continuity of the payoff functions to show that a strict equilibrium for one discount factor implies that nearby discount factors also gives an equilibrium.

Recall that in the proof of Proposition 2, $\underline{\delta}'$ was defined so that the contagious strategies with parameter $q'(\underline{\delta}')=1$ give a sequential equilibrium. An important intermediate step in the proof was to establish the existence of $\bar{\varepsilon}_1$ such that (9) and (10) held for all $\varepsilon < \bar{\varepsilon}_1$ and all $\delta \in [\underline{\delta}', 1)$. Substituting $\underline{\delta}'$ into these equations gives

$$\frac{\underline{\delta}'}{1-\underline{\delta}'} (f(1, \underline{\delta}', 1, \varepsilon) - f(2, \underline{\delta}', 1, \varepsilon)) > g + \eta/2$$

and

$$\frac{\underline{\delta}'}{1-\underline{\delta}'} (f(k, \underline{\delta}', 1, \varepsilon) - f(k+1, \underline{\delta}', \varepsilon)) < g - \eta/2 \quad \forall k \geq 2.$$

Restricting attention to values $\delta \in [\underline{\delta}', (1+\underline{\delta}')/2]$, we once again can easily establish bounds on the derivatives of the left-hand sides of the equations (11) and (12). For example, using expression (A3) from the proof of Proposition 2 we get

$$\begin{aligned} & \frac{\partial}{\partial \delta} \left(\frac{\delta}{1-\delta} (f(k, \delta, 1, \varepsilon) - f(k+1, \delta, 1, \varepsilon)) \right) \\ &= \sum_{t=0}^{\infty} (t+1) \delta^t (1+g) \text{Prob} \{o_t(t) \in C(t, k) \cap D(t) \cap \overline{E(t)}\} \\ &\cong \frac{1+g}{(1-\delta)^2} \\ &\cong \frac{4(1+g)}{(1-\underline{\delta}')^2}. \end{aligned}$$

Hence, we can find a value δ_1 such that for all $\delta \in [\underline{\delta}', \delta_1]$ and all $\varepsilon \leq \bar{\varepsilon}_1$ we have

$$\frac{\delta}{1-\delta} (f(1, \delta, 1, \varepsilon) - f(2, \delta, 1, \varepsilon)) > g + \eta/4 \quad (13)$$

and

$$\frac{\delta}{1-\delta} (f(k, \delta, 1, \varepsilon) - f(k+1, \delta, 1, \varepsilon)) < g - \eta/4 \quad \forall k \geq 2. \quad (14)$$

From here, the same steps as in the proof of Proposition 2 but with $\eta/2$ in place of η show that for sufficiently small ε , the strategies with $q=1$ give an equilibrium for all $\delta \in [\underline{\delta}', \delta_1]$. Now, the construction in Lemma 2 gives us an equilibrium without public randomizations for all $\delta \in [\underline{\delta}'/\delta_1, 1)$.

A further consequence of Lemma 2 is that the per period payoff to a player with discount factor δ of the no randomization equilibrium $s^{**}(\delta, \varepsilon)$ is *exactly* equal to the

TABLE II

Discount factors with and without public randomizations

		$g=0.01$	$g=1$	$g=10$
$M=4$	$\underline{\delta}$	0.03	0.68	0.95
	δ_1	0.03	1.00	1.00
	$\underline{\delta}/\delta_1$	0.96	0.68	0.95
$M=10$	$\underline{\delta}$	0.08	0.79	0.97
	δ_1	0.08	1.00	1.00
	$\underline{\delta}/\delta_1$	0.96	0.79	0.97
$M=100$	$\underline{\delta}$	0.35	0.89	0.985
	δ_1	0.36	1.00	1.00
	$\underline{\delta}/\delta_1$	0.96	0.89	0.985
$M=1000$	$\underline{\delta}$	0.57	0.93	0.990
	δ_1	0.60	1.00	1.00
	$\underline{\delta}/\delta_1$	0.96	0.93	0.990

per period payoff that the strategies with $q=1$ give a player with discount factor $\delta^{N(\delta)}, f(0, \delta^{N(\delta)}, 1, \epsilon)$. The function f is continuous in its second argument and $\delta^{N(\delta)} \rightarrow \underline{\delta}'$ as $\delta \rightarrow 1$, so for u_i being player i 's expected utility in the game with discount factor δ ,

$$\lim_{\epsilon \rightarrow 0} \lim_{\delta \rightarrow 1} u_i(s^{**}(\delta, \epsilon)) = \lim_{\epsilon \rightarrow 0} f(0, \underline{\delta}', 1, \epsilon).$$

This, however, is merely the limit of the expected payoff for a *fixed* discount factor as $\epsilon \rightarrow 0$ so efficiency in the limit is easy. For any $\gamma > 0$, we can simply choose T so that $(1 - \underline{\delta}')(1 + \underline{\delta}' + \dots + \underline{\delta}'^T) > 1 - \gamma/2$ then pick ϵ small enough so that with very high probability there are no ϵ -probability events in the first T periods, hence giving an expected payoff of at least $1 - \gamma$ in the game with ϵ noise. ||

If we had not worried about noise in this section, we could have found an equilibrium without public randomizations whenever $\delta \in [\underline{\delta}, \delta_1]$ where $\underline{\delta}$ is defined by

$$\underline{\delta}(1 - f(2, \underline{\delta}, 1)) = (1 - \underline{\delta})g$$

and δ_1 is defined either by

$$\delta_1(f(3, \delta_1, 1) - f(4, \delta_1, 1)) = (1 - \delta_1)g$$

or by $\delta_1 = 1$ if the equation above has no solution. Table II gives $\underline{\delta}$, δ_1 and $\underline{\delta}/\delta_1$ for a range of values of g and M . For $\delta \geq \underline{\delta}$ a cooperative sequential equilibrium exists with public randomizations, and for $\delta \geq \underline{\delta}/\delta_1$ one exists without them. Note that for many of the parameter values, δ_1 is in fact equal to one. In this case, eliminating public randomizations does not require any additional patience on the part of the players. When $g=0.01$, the difficulty in getting the players to carry out punishments results in much more patient play being necessary to support the equilibrium I have given.

5. CONCLUSION

In all of the results above, cooperation has been sustained in equilibrium by the use of "contagious" punishments which lead eventually to a breakdown of cooperation after a single deviation. The results illustrate the extent to which the convexity of the breakdown process can be exploited, and the interesting patterns of play which can arise in equilibrium. In addition, the contagious punishments are a fairly powerful tool for enforcing

cooperation. Besides the basic result that cooperation can be sustained despite a very large population of not unreasonably patient players with infrequent individual interactions, we have seen that cooperation is still possible with heterogeneity in time preferences or without public randomizations.

I have also argued that these results can be made far more robust than Kandori's first example suggests. Global stability is not a problem if public randomizations are available. In a limiting sense, the possibility that players may tremble and cheat accidentally is also not a problem even if players do not know the exact frequency of these trembles. All the attention paid to robustness in this sense should not be taken to necessarily imply that the paper is intended to be a plausible explanation for why cooperation has been observed in some particular situation. The equilibrium has a number of problems as a practical model: the cooperative equilibrium is at best one of a multiplicity of equilibria and when mistakes are introduced we may need extremely restrictive assumptions on the frequency of mistakes and the size of the population. The equilibrium is also very dependent on the assumption that all players are rational. I nonetheless think that the robustness is not merely of game theoretic interest, because it suggests that further analysis of properties other than the simplest notions of stability or limiting efficiency with noise is needed and may improve our understanding of behaviour in large populations.

Finally, I should note that I have also left one major question of game-theoretic interest unanswered. The results of this paper rely heavily on the fact that the prisoner's dilemma has a dominant strategy equilibrium. In light of Kandori's Folk Theorem for games with a more complex information structure, it would be interesting to know whether the results of this paper extend to a more general class of games. If so, we would have a much more general Folk Theorem. If not, we would have a sharper picture of the type of information transmission which is necessary to maintain cooperation.

APPENDIX

Proof of Proposition 2. I begin by establishing equations (7) and (8) which are analogous to equations (1) and (2) from the proof of Proposition 1. I will write $f(k, \delta, q, \varepsilon)$ for the per period continuation payoff of player i when at the start of period t , k players (including player i if $k > 0$) are playing according to phase II of the strategies described in the proof of Proposition 1. I wish to show that there exists $\delta' < 1$, $\eta > 0$ and a function $q' : [\delta', 1] \rightarrow [0, 1]$ such that (7) and (8) hold for all $\delta \in [\delta', 1]$.

Note first that because I have not yet introduced noise, $f(0, \delta, q'(\delta), 0) = 1$. I begin by establishing a degree of strict convexity of f . From equation (3) in the proof of Lemma 1 we know that

$$\begin{aligned} & ((f(1, \delta, q, 0) - f(2, \delta, q, 0)) - (f(2, \delta, q, 0) - f(3, \delta, q, 0))) \\ & = E_{\omega} [\sum_{t=0}^{\infty} (1 - \delta) q' \delta' (1 + g) I(o_1(t, \omega) \in (C(1, 1, \omega) - C(1, 2, \omega)) \cap D(t, \omega))]. \end{aligned}$$

The second term of this sum is

$$(1 - \delta) q \delta (1 + g) \text{Prob} \{o_1(t, \omega) \in (C(1, 1, \omega) - C(1, 2, \omega)) \cap D(1, \omega)\}. \quad (\text{A1})$$

If player 2 is matched with player M in period 0 under ω we have

$$\begin{aligned} 2 & \in C(1, 1, \omega) & M & \in C(1, 1, \omega) \\ 2 & \notin C(1, 2, \omega) & M & \notin C(1, 2, \omega) \\ D(1, \omega) & = \{2, M\}. \end{aligned}$$

Together, these imply

$$(C(1, 1, \omega) - C(1, 2, \omega)) \cap D(1, \omega) = \{2, M\}.$$

From this, we know that the probability term in (A1) is at least the probability that players 2 and M are matched in period 0 and that player 1 subsequently is matched against one of 2 or M in period 1. This probability is $2/(M-1)^2$.

Hence, for $\underline{\delta}$ as defined in Proposition 1, we have for any $\delta > \underline{\delta}$,

$$\frac{\delta}{1-\delta} ((f(1, \delta, 1, 0) - f(2, \delta, 1, 0)) - (f(2, \delta, 1, 0) - f(3, \delta, 1, 0))) \geq \frac{2\delta^2(1+g)}{(M-1)^2} \equiv \gamma. \tag{A2}$$

From equation (5) we know that

$$\frac{\delta}{1-\delta} (f(1, \delta, 1, 0) - f(2, \delta, 1, 0)) = g.$$

From expansion (6) in the proof of Proposition 1 it is immediate that

$$\frac{\partial}{\partial \delta} (f(1, \delta, 1, 0) - f(2, \delta, 1, 0))|_{\delta} > 0.$$

Thus for some $\eta < \gamma/2$ we can choose $\delta' \in (\underline{\delta}, 1)$ so that

$$\frac{\delta'}{1-\delta'} (f(1, \delta', 1, 0) - f(2, \delta', 1, 0)) = g + \eta.$$

By (A2) we know

$$\frac{\delta'}{1-\delta'} (f(2, \delta', 1, 0) - f(3, \delta', 1, 0)) < g - \eta.$$

Now, we simply set

$$q'(\delta) \equiv \delta'/\delta$$

and note from (6) that $\forall \delta \in [q'(\delta), 1)$

$$\frac{\delta q'(\delta)}{1-\delta} (f(k, \delta, q'(\delta), 0) - f(k+1, \delta, q'(\delta), 0)) = \frac{\delta'}{1-\delta'} (f(k, \delta', 1, 0) - f(k+1, \delta', 1, 0)).$$

As $q'(\delta) > q(\delta)$, players will not deviate in phase I of a model with no noise so

$$f(0, \delta, q'(\delta), 0) \geq f(1, \delta, q'(\delta), 0).$$

This establishes (7) and (8) as desired.

The next major step in the proof is to establish that the similar inequalities (9) and (10) hold for a model with sufficiently little noise. To do this, I extend expansion (3) to a model with noise. Note that

$$\begin{aligned} & f(k, \delta, q, \varepsilon) - f(k+1, \delta, q, \varepsilon) \\ &= E_{\omega} [\sum_{t=0}^{\infty} (1-\delta) q' \delta' (1+g) I(o_1(t, \omega) \in C(t, k, \omega) \cap D(t, \omega) \cap \overline{E(t, \omega)})] \end{aligned} \tag{A3}$$

where a realization of ω now includes also the set of players who “tremble” and play D accidentally in each period and $E(t, \omega)$ is defined to be the set of players affected by an ε -probability tremble up to and including time t . If $T(t, \omega)$ is the set of players who tremble at time t for a realization of ω , $E(t, \omega)$ can be formally defined by

$$E(0, \omega) = T(0, \omega)$$

$$E(t+1, \omega) = E(t, \omega) \cup T(t+1, \omega) \cup \{i | o_i(t, \omega) \in E(t, \omega)\}.$$

Using the expansions (3) and (A3) we get

$$\begin{aligned} & \frac{\delta q'(\delta)}{1-\delta} ((f(k, \delta, q'(\delta), 0) - f(k+1, \delta, q'(\delta), 0)) - (f(k, \delta, q'(\delta), \varepsilon) - f(k+1, \delta, q'(\delta), \varepsilon))) \\ &= E_{\omega} [\sum_{t=0}^{\infty} q'(\delta)^{t+1} \delta'^{t+1} (1+g) I(o_1(t, \omega) \in C(t, k, \omega) \cap D(t, \omega) \cap E(t, \omega))] \\ &\leq E_{\omega} [\sum_{t=0}^{\infty} \delta'^{t+1} (1+g) I(o_1(t, \omega) \in E(t, \omega))]. \end{aligned} \tag{A4}$$

Given $\eta > 0$ as defined above, we can choose T such that

$$\frac{\delta'^T}{1 - \delta'} < \frac{\eta}{4(1+g)}.$$

Next, choose $\bar{\varepsilon}_1$ sufficiently small such that

$$\text{Prob} \{E(T, \omega) \neq \emptyset\} < \frac{\eta}{4(1+g)(1 + \dots + \delta'^T)}.$$

Now, for any $\delta \in [\delta', 1)$ and any $\varepsilon < \bar{\varepsilon}_1$, the right-hand side of equation (A4) is bounded above by $\eta/2$. This and equations (7) and (8) gives

$$\frac{\delta q'(\delta)}{1 - \delta} (f(1, \delta, q'(\delta), \varepsilon) - f(2, \delta, q'(\delta), \varepsilon)) > g + \eta/2$$

and

$$\frac{\delta q'(\delta)}{1 - \delta} (f(2, \delta, q'(\delta), \varepsilon) - f(3, \delta, q'(\delta), \varepsilon)) < g - \eta/2.$$

The first equation is (9). Using the expansion (A3) in place of (3) it is easy to see that the result of Lemma 1 carries over to the model with ε -noise. This and the second equation above gives us (10).

Now that (9) and (10) have been established, I will proceed to show that there are no profitable deviations from either phase I or phase II play in the ε -constrained game. The phase II case is easier so I'll start with that. Note that we can rewrite (10) to give $\forall \varepsilon < \bar{\varepsilon}_1$, $\delta \in [\delta', 1)$, and $k \geq 2$,

$$\delta q'(\delta) (f(k, \delta, q'(\delta), \varepsilon) - f(k+1, \delta, q'(\delta), \varepsilon)) < (1 - \delta)g.$$

As in the proof of Proposition 1, the right hand side of this expression is the short term loss when a player plays C instead of D in phase II and is matched with someone who plays C . The expectation over k of the left-hand side is the expected future gain. Clearly, the future gain is too small to make a deviation profitable.

The discussion of phase I play is more complicated than before because a player in phase I must assign probability $r_k > 0$ to the event that unbeknown to him, k other players are already playing according to phase II or will tremble and play D in the current period. Keep in mind that r_k is a function both of ε and of the history of the game. I will show, however, that for ε sufficiently small this uncertainty is small regardless of the history of the game.

To show that player 1's best response whenever he is in phase I in period t is to play C , I will *not* show directly that his expected payoff from playing C in period t and then following his equilibrium strategy is better than his expected payoff from playing D in period t then following his equilibrium strategy. Instead, I will compare the payoff from playing C in period t then switching to phase II play in period $t+1$ to the payoff from playing D in period t and continuing according to phase II. Player 1's period t action has no effect on play after any period $t+s$ in which $q_{t+s} > q'(\delta)$. We have already seen that playing D in phase II is a best response so that the latter strategy gives the greatest possible expected payoff to a player who plays D in period t . If the former is greater, the best response must involve playing C in period t .

To compare the payoffs of the two strategies, look first at the period- t outcome. If player 1 plays D in period t he gains g whenever $o_1(t)$ plays C and avoids a loss of l whenever $o_1(t)$ plays D . Hence the short term gain is

$$\sum_{k=0}^{M-1} r_k \left(\frac{M-k-1}{M-1} g + \frac{k}{M-1} l \right).$$

In the future, a player who plays D in period t can never be better off because both strategies prescribe the same play from period $t+1$ on and there are always either the same number or more players in phase II in period $t+1$. When $k=0$ and there are also no ε -probability trembles in period $t+1$, the player who plays D in period t is worse off, obtaining a continuation payoff of $f(2, \delta, q'(\delta), \varepsilon)$ instead of $f(1, \delta, q'(\delta), \varepsilon)$. The discounted expected loss is then at least

$$r_0(1 - \varepsilon)^{M-1} \frac{\delta q'(\delta)}{1 - \delta} (f(1, \delta, q'(\delta), \varepsilon) - f(2, \delta, q'(\delta), \varepsilon)).$$

To show that playing C is better in period t it thus suffices to show that this loss outweighs the short-term gain. Using (9), it will suffice to show

$$r_0(1 - \varepsilon)^{M-1}(g + \eta/2) \geq r_0g + (1 - r_0) \max(g, l).$$

We can choose $\bar{\varepsilon}_2$ such that

$$(1 - \varepsilon)^{M-1}(g + \eta/2) \geq g + \eta/4$$

for all $\varepsilon < \bar{\varepsilon}_2$. It then only remains to establish

$$r_0\eta/4 \geq (1 - r_0) \max(g, l) \tag{A5}$$

for ε sufficiently small.

At first look one might think that if the game has been going on long enough then player 1 will be fairly sure that someone must have trembled. This reasoning suggests that the r_0 term might not dominate in (A5). However, it is important to keep in mind that r_0 is not an unconditional probability, but rather the conditional probability that no one has trembled since the last time s that $q_s > q'(\delta)$ occurred given that no opponent of player 1 has played D since that time. Write $r_0(t)$ for the probability that all other players are cooperating conditional on it having been t periods since $q_s > q'(\delta)$ and on cooperation only having been observed in that time. To show in fact that

$$\liminf_{\varepsilon \rightarrow 0} r_0(t) = 1 \tag{A6}$$

take any $\zeta > 0$. We can choose T_1 so that

$$\text{Prob}\{\text{Player 1 is still in phase I} \mid \text{Some player was in phase II } T_1 \text{ periods ago}\} < \frac{(1 - \zeta)(1 - \zeta/2)}{2}.$$

Next choose $\bar{\varepsilon}_3$ so that

$$(1 - \bar{\varepsilon}_3)^{T_1 M} > 1 - \zeta/2$$

We now show by induction that for any $\varepsilon < \bar{\varepsilon}_3$ and any t we have $r_0(t) > 1 - \zeta$. First, for $t \leq T_1$, the probability that there has been no tremble is at least $1 - \zeta/2$ and conditioning on not seeing a tremble only increases this probability. Next, if for some $T \geq T_1$ we have $r_0(t) > 1 - \zeta$ for all $t \leq T$, then $1 - r_0(T+1)$ is less than the sum of the probability that some player was in phase II in period $T+1 - T_1$ conditional on player 1 still being in phase I, and the probability that there has been a tremble in the last T_1 periods. The first probability by Bayes' rule is less than

$$\frac{(1 - r_0(T+1 - T_1))(1 - \zeta)(1 - \zeta/2)/2}{(1 - r_0(T+1 - T_1))(1 - \zeta)(1 - \zeta/2)/2 + r_0(T+1 - T_1)(1 - \zeta/2)} < \frac{\zeta(1 - \zeta)(1 - \zeta/2)}{2(1 - \zeta)(1 - \zeta/2)} = \frac{\zeta}{2}.$$

The second term is at most $\zeta/2$. Hence, by induction $r_0(t) > 1 - \zeta$ for all t . Choosing $\bar{\varepsilon}$ smaller than $\bar{\varepsilon}_1$, $\bar{\varepsilon}_2$, and $\bar{\varepsilon}_3$ we get the sufficient condition (A5) for no deviations in phase I. This concludes the proof that $s^*(\delta, \varepsilon)$ is a sequential equilibrium of the ε -constrained game.

Finally, the proof of 3, that we get efficiency in the limit, is relatively easy. Consider the largest possible effect that a single tremble by player j in t period t can have on player i 's total payoff anywhere on the path of the equilibrium with noise. This tremble can only affect player i 's payoff in period t and in any future period until the first time $q_{t+s} < q'(\delta)$. Thus, the expected loss caused by this single tremble is at most

$$\delta' \sum_{s=0}^{\infty} (1 - \delta) \delta^s q'(\delta)^s (1 + g + l) = (1 + g + l) \frac{\delta'(1 - \delta)}{1 - \delta'}.$$

Player i 's expected per period payoff is equal to 1 minus the expected loss from each possible tremble times the probability of that tremble occurring. This gives

$$\begin{aligned} f(0, \delta, q'(\delta), \varepsilon) &\geq 1 - (1 - \delta) \sum_{t=0}^{\infty} (1 + g + l) \frac{\delta'}{1 - \delta'} M \varepsilon \\ &= 1 - \frac{(1 + g + l) M \varepsilon}{1 - \delta'} \end{aligned}$$

Clearly

$$\lim_{\varepsilon \rightarrow 0} \lim_{\delta \rightarrow 1} f(0, \delta, q'(\delta), \varepsilon) = 1. \quad \parallel$$

Acknowledgements. I would like to thank Sara Fisher Ellison, Lones Smith, and three referees for their comments and Drew Fudenberg for comments, suggestions, and guidance at all stages of this project. Financial support from the Sloan Foundation and from National Science Foundation grant SBR-9310009 is acknowledged.

REFERENCES

- AUMANN, R and SHAPLEY, L. (1976), "Long Term Competition—A Game Theoretic Analysis" (Mimeo, Hebrew University).
- ELLISON, G. (1993), "Learning, Local Interaction and Cooperation", *Econometrica*, **61**, 1047–1071.
- FRIEDMAN, J. (1971), "A Non-cooperative Equilibrium for Supergames", *Review of Economic Studies*, **38**, 1–12.
- FUDENBERG, D. and MASKIN, E. (1986), "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information", *Econometrica*, **54**, 533–554.
- FUDENBERG, D. and MASKIN, E. (1991), "On the Dispensibility of Public Randomizations in Discounted Repeated Games", *Journal of Economic Theory*, **53**, 428–438.
- FUDENBERG, D., LEVINE, D. and MASKIN, E. (1993), "The Folk Theorem with Imperfect Public Information" (Mimeo, Harvard University).
- GREEN, E. (1980), "Non-cooperative Price Taking in Large Dynamic Markets", *Journal of Economic Theory*, **22**, 155–182.
- GREIF, A. (1989), "Reputation and Coalitions in Medieval Trade: Evidence on the Maghribi Traders", *Journal of Economic History*, **69**, 857–882.
- HARRINGTON, J. (1991), "Cooperation in a One-shot Prisoners' Dilemma" (Mimeo, Johns Hopkins).
- KANDORI, M. (1992), "Social Norms and Community Enforcement", *Review of Economic Studies*, **59**, 63–80.
- MILGROM, P., NORTH, D. and WEINGAST, B. (1990), "The Role of Institutions in the Revival of Trade: The Law Merchant, Private Judges, and the Champagne Fairs", *Economics and Politics*, **2**, 1–23.
- OKUNO-FUJIWARA, M. and POSTLEWAITE, A. (1990), "Social Norms and Random Matching Games" (CARESS Working Paper No. 90-18, University of Pennsylvania).
- ROSENTHAL, R. (1979), "Sequences of Games with Varying Opponents", *Econometrica*, **47**, 1353–1366.
- SABOURIAN, H. (1990), "Anonymous Repeated Games with a Large Number of Players and Random Outcomes", *Journal of Economic Theory*, **51**, 92–110.