Community enforcement using modal actions*

Sam Jindani

Department of Economics, University of Oxford, United Kingdom

16 November 2019

Abstract

Can cooperation be sustained in large populations? This paper studies settings in which a large group of players is rematched at random each period. In such settings cooperation cannot be sustained by an equilibrium unless deviators are sanctioned by third parties. This is known as the problem of *community enforcement*. Previous analyses have relied on strong assumptions about what information players have access to. This paper shows that when players are matched with multiple partners in each period, it is possible to limit the amount of information required to support cooperative outcomes. The results hold for general games and for equilibria that are robust to noise.

Keywords: repeated games, random matching, community enforcement, information transmission, modal actions

1. Introduction

The folk theorem for repeated games can be seen as a formal solution to the problem of cooperation between individuals.[1] For instance, it implies that players in a repeated prisoner's dilemma can cooperate, despite incentives to cheat one another, by threatening credible sanctions if one of them deviates. More generally, it implies that in many repeated games efficient outcomes can be enforced by an equilibrium if players are sufficiently patient. The canonical folk theorem applies to settings in which players interact with the

*I thank my supervisors, Peyton Young and Tom Norman, as well as Alan Beggs, Yuval Heller, Ed Hopkins and Meg Meyer for their insightful comments. I also thank the referees and the associate editor for their extremely constructive suggestions. I gratefully acknowledge financial support from the Economic and Social Research Council (award number ES/J500112/1) and the University of Oxford Department of Economics.

Email address: `sam.jindani@economics.ox.ac.uk` (Sam Jindani)

[1] Aumann and Shapley 1976; Rubinstein 1979; Fudenberg and Maskin 1986.

same partners every period. However in many economic settings, agents interact with different partners at different times and interactions with any given partner may be too infrequent for the standard folk theorem to apply. Can cooperation still be sustained in such settings?

Following Kandori (1992) and Okuno-Fujiwara and Postlewaite (1995), I model this situation as a *repeated game with random matching*. Each period, players are randomly matched to play a stage game. Because any two players meet infrequently, sanctions against players who deviate must be carried out by third parties. This creates a problem of information transmission: If a player deviates, how will her partners in future periods know that she should be sanctioned? And if one of her future partners fails to sanction her, how will the partner's future partners know that she, in turn, should be sanctioned? And so on. This is what Kandori termed the problem of *community enforcement*. In a standard repeated game, the problem does not arise, because players observe every action taken by every player; in a repeated game with random matching, this would be implausible. The question is how much information transmission is necessary to sustain cooperation.

Kandori's seminal paper presents two approaches. First, he shows that cooperation can be sustained in the prisoner's dilemma without information transmission, but at the price of instability: in his equilibrium, players cooperate as long everyone they encounter cooperates; but as soon as one of their opponents defects against them, they defect in all future periods. This is known as a *contagion equilibrium*, because defection spreads from player to player through the population. This type of equilibrium is only feasible in certain stage games such as the prisoner's dilemma, where sanctions correspond to a stage-game Nash equilibrium. Moreover, the equilibrium, as Kandori notes, is not robust to noise, in the sense that if a single player makes a mistake and defects, the whole population starts to defect and never returns to cooperation.[2]

Kandori's second approach relies on strong assumptions about information transmission. Kandori assumes that there is an exogenous information clearinghouse that recognises deviations and assigns publicly observable 'labels' to players who deviate. Given this assumption, he is able to prove a robust folk theorem for (almost) general games and equilibria that are robust to noise. In his equilibrium, the labels denote the number of periods a player should be sanctioned for. If a player deviates, her label is updated to a positive integer. Players who are matched with her in the ensuing periods

[2]Ellison (1994) shows that returning to cooperation can be achieved, although this requires public randomisations (see the literature review in section 2).

observe this label and know that they are supposed to sanction her; if they fail to do so their label is updated and they will be sanctioned in turn. In his conclusion, Kandori notes: 'Perhaps the most important question which is unanswered by the present paper concerns the way in which the information transmission postulated in our model is implemented' (Kandori 1992, p. 77). Okuno-Fujiwara and Postlewaite make a similar assumption in their paper (with 'statuses' instead of 'labels').

Most of the literature on repeated games with random matching has followed Kandori's first approach in assuming little or no information transmission between players. These results tend only to apply to games like the prisoner's dilemma or to be limited to non-robust or weak equilibria (see the literature review in section 2). I take the second approach. I show that it is possible to obtain a folk theorem for general games and for strict and robust equilibria under significantly weaker assumptions about information transmission than those made by Kandori.

The results are obtained using a simple construction: In a standard repeated game with random matching, players are paired with a single partner each period. I assume that players are matched with multiple partners in each period, and play the same stage game with each. This drastically changes the nature of the game, because it becomes easier to identify deviators. In a standard repeated matching game without multiple matching, players may require a large amount of information to determine whether a partner deviated in a previous period. For instance, in the prisoner's dilemma a player who defected in a given period may have deviated, or may have been sanctioning a player who deviated. To find out, one needs to determine whether the player she was matched with had deviated previously, which may require determining whether the players that player was matched with in previous periods deviated, and so on. In general, the amount of information required increases over time and is unbounded. In contrast, when players are matched with multiple partners, one can determine whether or not a player deviated in a given period just by looking at her actions and the actions of other players in that period – specifically, it suffices to check whether a player sanctioned someone when others did not, or failed to sanction someone when others did. That is, one can compare a player's actions with the corresponding *modal actions* in that period. If they differ, then one can infer that the player deviated.[3]

[3] It is worth noting that there exist bounded-memory folk theorems for standard repeated games (*i.e.* games without random matching) in which the amount of information required by players does not increase over time. See for instance Barlo et al. 2016. However the strategies involved are complex, requiring players to set aside periods to play

I assume that players can observe partners' actions in recent periods as well as the corresponding modal actions. I show that this is sufficient to obtain a folk theorem for general games. In the equilibrium, players sanction partners who were non-modal in the recent past. Then it is in each player's interest not to deviate on the equilibrium path, because a player who does so will be non-modal and will be sanctioned. It is also in each player's interest to sanction a player who deviates, because failing to do so would mean failing to sanction someone whom others sanction, which would trigger sanctions in the following periods. In this framework, the amount of information required by players is uniformly bounded from above across all possible histories. In particular, unlike in Kandori's framework, the amount of information required does not increase over time or in the size of the population; moreover, the result does not rely on an exogenous information clearinghouse.

The exact amount of information each player requires depends on how far in the past they have to look back, which in turn depends on how long deviators have to be sanctioned for. I argue that for relevant applications, this may not be very long at all. For instance, in a worked example using the prisoner's dilemma, I show that players only need to look back one period to sustain cooperation.[4]

The model can be thought of as applying to the following setting: Consider a relatively large and tight-knit community such as a medium-sized town. Each townsperson interacts with a number of others on a regular basis. However, any given pair of townspeople may interact infrequently.[5] The interactions between townspeople could be economic transactions, such as buying and selling goods and services, or purely social interactions, ranging from stopping to exchange pleasantries, to inviting someone to a social function, to providing assistance when necessary.

There may be a number of different social norms in place in the town. For instance, townspeople could be expected not to be rude or disrespectful to one another; or they might be expected to pay their debts and return favours. These norms are enforced by the community. Depending on the severity of the violation, the sanction may be more or less severe. For instance, suppose a townsperson is rude towards someone at a social function. Then people may

---

'signalling sequences' at regular intervals to identify deviators. Moreover, it is not clear that the results would extend to the random matching case. I discuss this more in section 6.

[4]The prisoner's dilemma is, of course, a particularly important application of the canonical folk theorem. It can be used to model trade or any number of situations in which players must trust each other to cooperate.

[5]One could also consider a neighbourhood association, a church, or a year group at a university or school.

stop inviting her to other events. If someone fails to pay a debt, people may refuse to do business with her. More severe violations could lead to complete ostracism. Crucially, deviations are sanctioned by the whole community. Being sanctioned just by the person who was harmed by an agent's deviation would be ineffective.

The problem of community enforcement is the specific mechanics of how these sanctions are carried out by members of the community. In particular, a key issue is how information about who should be sanctioned is transmitted to relevant agents. In this setting, there is no central clearinghouse that would label deviators as in Kandori's framework. It would also be unrealistic to assume all townspeople have full information about all other agents' past actions since the beginning of the game. For one thing, the population is too large; for another, it is unclear when the game even began.

The present paper offers an answer to this problem. In the setting described above, it is natural to assume that agents are matched with multiple other agents in each period. In that case, the paper shows that players' statuses as deviators or non-deviators can be tied to whether or not they were modal. This considerably simplifies the information transmission required. The only thing agents need to know is whether a partner was non-modal in the recent past; that is, the townspeople need to be able to find out if a person they interact with acted in a way that other townspeople did not. This kind of information is readily available to different members of the community and can be shared and verified easily – one can imagine the townspeople gossiping and sharing information about a particular agent's violation of a norm. Multiple matching provides one key feature: what an agent should have done in any given interaction is clear and easy to identify.

Like Kandori's labelling equilibrium, the proposed equilibrium is robust to noise. In a real-world setting, such as the one described above, noise is pervasive. Players might fail to sanction someone when they were meant to sanction them, or sanction someone they weren't meant to. They could observe their partners actions inaccurately, or they might simply be confused. Some players could even be irrational and have some other basis for picking actions than best-responding. Given this, robustness to noise is an important property of an equilibrium. I consider two different definitions of robustness and show that the proposed equilibrium is robust under both. First, Kandori defines a notion of robustness he calls *global stability*, according to which players' continuation payoffs must converge to their equilibrium-path values from any history. In other words, the effect of any individual deviation must eventually die out. This is a relatively straightforward and intuitive definition of robustness. It is satisfied by Kandori's labelling equilibrium but not his contagion equilibrium. The equilibrium of the present paper is straightforwardly globally stable, since punishment phases are finite.

Second, Ellison (1994) argues that if noise occurs in every period, global stability may not be an appropriate definition of robustness. If noise is persistent, an equilibrium could be globally stable, yet if it takes too long for the impact of an error to die out players could have expected payoffs far from their values in the case without noise. Ellison models this setting by assuming that each period each player *trembles* with some small probability and plays an action other than the one she intended to play. He shows that contagion equilibria in the prisoner's dilemma can be made robust to noise in the sense that, as noise tends to zero, the strategies remain an equilibrium and the players' continuation payoffs converge to the equilibrium-path payoffs. Although he requires public randomisations to obtain global stability, this is not required for robustness to persistent noise.

In the spirit of Ellison, I show that the proposed equilibrium is robust to persistent noise. As in Ellison's analysis, I assume players tremble with some small probability. However, I also consider two additional types of noise: observation errors, whereby players observe their partners' past actions inaccurately, and noisy players, whereby some players always choose their actions at random. Given the importance of information transmission in random matching games, robustness to observation errors is a particularly desirable form of robustness. Robustness to noisy players is also significant, because, as Ellison notes, in a contagion equilibrium a single noisy player can cause defection throughout the population in every period. I show that the proposed equilibrium is robust to trembles and noisy players. The equilibrium is not robust to observation errors when they are the only source of noise. However it is robust when there is also another type of noise and the probability of an observation error is small relative to the other source of noise.[6]

The rest of this paper is organised as follows: In section 2, I provide an overview of the literature on community enforcement. In section 3, I set up the model of repeated matching games with multiple partners. In section 4, I prove the main result and solve a worked example with the prisoner's dilemma. Finally, in section 5, I consider robustness to noise.

[6]Observation errors are a type of imperfect private monitoring. The reason the equilibrium breaks down when they are the only source of noise is that players always attribute non-modal actions to observation errors rather than deviations, and since observation errors are independent they assume other players will not have observed the non-modal action. Therefore deviations are not sanctioned and players can deviate with impunity.

2. Related literature

As discussed in the preceding section, the seminal papers in the literature are Kandori 1992 and Okuno-Fujiwara and Postlewaite 1995.[7] Another early contribution is due to Harrington (1995), who considers a model in which time is continuous and players meet at random intervals.

As discussed above, Ellison (1994) shows that contagion equilibria in the prisoner's dilemma can be made robust to noise, using a public randomisation device. As in Kandori's contagion equilibrium, players cooperate on the equilibrium path and start defecting as soon as they are matched with a partner who defects. However, players return to cooperating as soon as the public randomisation device takes a certain value. The value is chosen to set the expected duration of the defection phase, which ensures that the strategies constitute an equilibrium. This equilibrium satisfies global stability, since play eventually returns to cooperation with probability one. As discussed above, it is also robust to persistent noise. Ellison also proves a result for the case in which there are no public randomisations. He constructs an equilibrium in which defections are spread out indefinitely over time, between periods of cooperation. (He says that we might see 'players cheating on every third Friday but cooperating on all other days' [p. 581].) This equilibrium is robust to persistent noise but is not globally stable.

Takahashi (2010) considers the prisoner's dilemma and assumes that players can observe their own past actions and their partners' past actions. The equilibria he constructs are related to concepts in the imperfect private monitoring literature, and in particular to belief-free equilibria. As in that literature, the strategies involved are complex and players are always indifferent between sanctioning and not sanctioning following a deviation. His equilibria are not guaranteed to be globally stable.

Deb (2019), also drawing on the imperfect private monitoring literature, obtains a folk theorem for general games. She assumes matched players can engage in cheap talk before each stage game, which allows them to identify one another. Since information is cheap talk, players prove their identity by playing 'signatures' that depend on their and their partner's last interaction. As with Takahashi's equilibria, the strategies are complex and involve players being indifferent.

Deb and González-Díaz (2019) have strict equilibria for games that satisfy the one-sided incentive problem (a class that includes the prisoner's dilemma). They assume no information transmission. The equilibria are similar to Kandori's and Ellison's contagion equilibria, but with an added

[7] Sugden (1986, chap. 7) provides a simpler version of Kandori's second result.

'trust-building' phase at the beginning of the game. The equilibria are not globally stable, although the authors conjecture that they could be made robust to persistent noise as considered by Ellison.

Dilmé (2016) looks at games in which players can help one another at a cost. The cost of helping is lower than the benefit created, so that helping is always socially beneficial. He considers the case where there is a small proportion of irrational players, and constructs an equilibrium that is robust to these players. The equilibrium is belief-free, so that players are always indifferent. When the population is large, the strategies used are close to tit-for-tat.

Heller and Mohlin (2017) also look at a setting where there is a small proportion of irrational players, but in the prisoner's dilemma. In each stage game, players observe a sample of their opponent's past actions. (The authors also have in mind a setting in which information transmission occurs through gossip or word of mouth, and in which the beginning of the game is uncertain.) They show that if the incentive to defect is too large, cooperation cannot be sustained. If the incentive to deviate is not too large, there is an essentially unique cooperative equilibrium in which some players defect when they observe two or more defections, and some when they observe one or more.[8]

3. Repeated matching games with multiple partners

The setup of the model broadly follows Kandori 1992, except that players are matched with multiple partners each period.

There is an even number of players, represented by the set $N := \{1, 2, \ldots, 2n\}$, where $n \geq 3$. The population is divided into two subpopulations $N_1 := \{1, 2, \ldots, n\}$ and $N_2 := \{n+1, n+2, \ldots, 2n\}$. I will refer to players in N_1 as type-1 players and to players in N_2 as type-2 players.[9] In each period each type-1 player is matched with s different type-2 players, where $s \geq 3$ is a factor of n, to play a stage game $\mathscr{G}$.[10]

[8]Of course, there are other approaches to studying cooperation in large populations. The literature outlined above focuses on best-response equilibria between perfectly rational agents. In contrast, evolutionary game theory studies interactions between imperfectly rational agents, and focuses on dynamics or steady states. A seminal article in that literature is Axelrod and Hamilton 1981. For an overview of evolutionary game theory see Weibull 1995 or Sandholm 2010.

[9]We could also have a model in which there were no subpopulations and players were randomly assigned a type each period in addition to being assigned partners.

[10]Note the assumption that interactions happen simultaneously, or at least simultaneously enough that players cannot observe what happened in other interactions of the same

Let A_1 and A_2 be the action sets of type-1 and type-2 players in $\mathscr{G}$, respectively, and let g_1 and g_2 be their payoff functions. Suppose A_1 and A_2 are finite. Let $A \coloneqq A_1 \times A_2$, with typical element a. Let V denote the payoff space of $\mathscr{G}$; that is, V is the convex hull of

$$\{v \in \mathbb{R}^2 : v = (g_1(a), g_2(a)) \text{ for some } a \in A\}.$$

As in Kandori's model, I rule out mixed actions. The only way in which this restricts the results is that, in general, the minmax payoffs without mixed actions may be higher than the minmax payoffs when mixed actions are allowed. It is straightforward to show that the results still hold when one allows mixed actions, provided one also allows either observability of mixed actions or public randomisations.

Each period, matches are realised according to a stochastic rule. Let $\mu(i,t)$ be the set of player i's realised matches in period t (so if i is of type 1, $\mu(i,t)$ is a subset of the set of type-2 players). Kandori assumes a general matching rule. This creates an incentive to deviate for players who would expect to sanction their partners for a number of periods. To address this, Kandori assumes that the stage game is such that sanctioners can be rewarded for carrying out sanctions. However in the equilibrium below, players will pick specific actions in the sanctions phase that make Kandori's approach infeasible. Therefore, instead of allowing a general matching rule I assume *uniform random matching* (the probability of any two matches in a given period is equal, and matches are independent across periods). I also assume the population is large relative to the maximum length of sanctions. Together these ensure that, in expectation, players will not be matched with too many deviators in any subgame. Kandori's assumption about the stage game is then not necessary, so the results hold for general games.

Players evaluate infinite sequences of payoffs using the discounting criterion with shared discount factor $\delta \in (0,1)$. Because players receive s stage-game payoffs per period, I define repeated-game payoffs to be discounted sums of *average* stage-game payoffs. Specifically, suppose player i's average stage-game payoff in each period t is v_i^t. Player i's average discounted payoff

---

period until the period is over. This simplifies analysis considerably. Although it is an approximation, it seems natural to assume that information about a particular interaction will take some time to reach other players, so that the approximation may be close enough. Note also that the requirement that s be a factor of n guarantees that matchings are possible.

for the sequences of average payoffs $\mathbf{v}_i = (v_i^t)_{t=1}^{\infty}$ is

$$u_i(\mathbf{v}_i) := (1-\delta)\sum_{t=1}^{\infty} \delta^{t-1} v_i^t.$$

Since each v_i^t is an average, we have $u_i(\mathbf{v}_i) \in V$. Let $\underline{v}_1$ be the minmax payoff for type-1 players and let m_2 be one of the corresponding minmaxing actions of type-2 players. Formally,

$$\underline{v}_1 := \min_{a_2} \max_{a_1} g_1(a_1, a_2), \text{ and}$$
$$m_2 \in \arg\min_{a_2} \max_{a_1} g_1(a_1, a_2).$$

Define $\underline{v}_2$ and m_1 analogously. The set of feasible and strictly individually rational payoff pairs is

$$V^* := \{v \in V : v \gg (\underline{v}_1, \underline{v}_2)\}.$$

Assume V^* is non-empty. Finally, it will be convenient to define the maximum payoffs $\bar{v}_1 := \max_a g_1(a)$ and $\bar{v}_2 := \max_a g_2(a)$.

This repeated matching game represents situations in which agents interact with different partners at different times. In the example of section 1, interactions between townspeople could be modelled by the generic prisoner's dilemma in figure 1. In each interaction, the players can either help one another and cooperate, or defect and harm their partners. One could also imagine that interactions between townspeople could be modelled by other stage games. The results below will hold for any stage game, including the prisoner's dilemma.

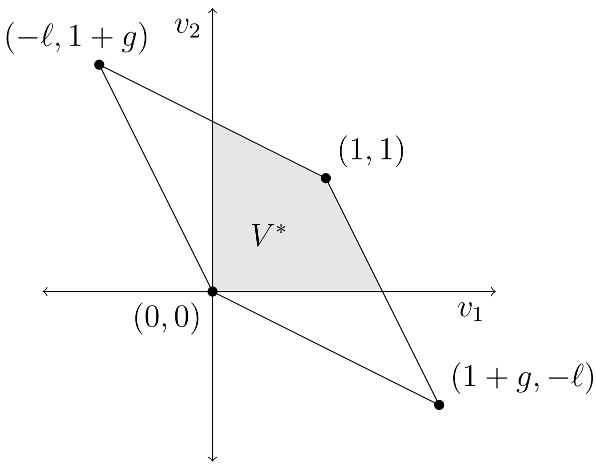
If the repeated matching game described above were a game with perfect information, it would be straightforward to adapt Fudenberg and Maskin's (1986) folk theorem to apply. However, since players are unlikely to be able to observe all actions in a large population, it would seem reasonable to restrict players' information. The question is how much information is necessary to sustain a folk theorem.

4. Observable modal actions

Before turning to the result, I introduce some notation.

If i was matched with j in period t, let a_{ij}^t be the action chosen by i in the stage game with j at t. Let $\boldsymbol{\alpha}_i^t$ be the s-tuple of actions taken by i at t; that is,

$$\boldsymbol{\alpha}_i^t := (a_{ij}^t)_{j \in \mu(i,t)}.$$



	C	D
C	$1, 1$	$-\ell, 1+g$
D	$1+g, -\ell$	$0, 0$

Figure 1: Game matrix and payoff space for the generic prisoner's dilemma; $g, \ell > 0$, and $g < 1 + \ell$.

Let θ_i^t be the set of the most frequent actions played against i at t. For example, suppose the stage game is the prisoner's dilemma of figure 1 and suppose $s = 4$. If three of the players i was partnered with at t picked C against i and one picked D, then $\theta_i^t = \{C\}$. If two of the players picked C against i and two picked D, then $\theta_i^t = \{C, D\}$. I will refer to θ_i^t as the set of *modal actions* against i at t.

Denote the s-tuple of modal sets corresponding to $\boldsymbol{\alpha}_i^t$ by $\boldsymbol{\vartheta}_i^t$; that is

$$\boldsymbol{\vartheta}_i^t := (\theta_j^t)_{j \in \mu(i,t)}.$$

In the equilibrium, players will compare $\boldsymbol{\alpha}_i^t$ to $\boldsymbol{\vartheta}_i^t$ to determine whether i deviated at t. Specifically:

Definition 1. Player i was *non-modal* at t if there exists some j such that $j \in \mu(i,t)$ and $a_{ij}^t \notin \theta_j^t$. Otherwise, i was *modal* at t.

Then, in particular, a player who failed to sanction someone whom others sanctioned will be non-modal; so will a player who sanctioned someone whom others did not sanction. Note that determining whether or not some player i was modal at t requires knowing exactly $\boldsymbol{\alpha}_i^t$ and $\boldsymbol{\vartheta}_i^t$. One can think of $\boldsymbol{\vartheta}_i^t$ as specifying the actions i should have played in order to be modal.[11]

[11] Provided i's actions cannot affect the corresponding model actions, which will be the case in equilibrium.

I assume that players have information about their partners' actions in the recent past as well as the corresponding modal actions. This ensures that players can determine whether or not their partners were modal. To be precise, in period t player i knows who she is matched with and the actions she and each current partner took in the previous p periods; that is, she knows $\mu(i,t)$ and $(\boldsymbol{\alpha}_j^\tau)_{\tau=t-1}^{t-p}$ for each $j \in \mu(i,t)$ and for $j = i$. Player i also knows, for each current partner and herself, the modal actions from the previous p periods corresponding to her or the partner's actions; that is, $(\boldsymbol{\vartheta}_j^\tau)_{\tau=t-1}^{t-p}$ for each $j \in \mu(i,t)$ and for $j = i$. She has no other information. I will refer to this as a *repeated matching game with multiple matching and observable modal actions*. The integer p can be thought of as the players' memory. I assume the population size n is large relative to p. Then, given uniform random matching and appropriate beliefs, in expectation, players will not be matched with too many deviators in any subgame.

Suppose we want players to receive some payoff pair $v \in V^*$. I construct an equilibrium in which players sanction those who were non-modal in previous periods. This mechanism not only disincentivises players from deviating on the equilibrium path, but also incentivises them to sanction those who deviate. A player who deviates on the equilibrium path will be playing a non-modal action, which means that she will be sanctioned in ensuing periods. A player who fails to sanction someone who played a non-modal action will herself be playing a non-modal action, so she will be sanctioned too.

Proposition 1. Every point $v \in V^*$ in the repeated matching game with multiple matching and observable modal actions is sustained by a strict, sequential equilibrium if δ and p are large enough, and n is large enough relative to p.

Proof. Assume there exists some $a \in A$ that achieves the designated payoff pair v. If there is no such action profile one can assume public randomisations are available or appeal to the techniques in Fudenberg and Maskin 1991.

The idea is that a player matched with someone who was non-modal should sanction her by playing the minmax action m_ℓ. In games where neither player was non-modal or where the sanctions phase is over, players pick a_ℓ. The proof is similar to the proof of theorem 1 in Fudenberg and Maskin 1986, except that unlike in Fudenberg and Maskin, players do not mutually minmax each other during the sanctions phase. This is because mutual minmaxing would make the sanctioned player non-modal against partners who were not being sanctioned. That is, if player i were supposed to be sanctioned, she would pick m_ℓ against her partners, but the modal action against her partners would normally be a_ℓ, so i would be non-modal.

Instead I suppose that the sanctioned player picks a_ℓ as normal (unless the other player also deviated).[12]

Suppose players i and j are matched in period t. Let ℓ be i's type. Fix an integer $q \leq p$ to be the length of sanctions. The strategy is as follows:

> If $t = 1$, player i picks a_ℓ. If $t > 1$ and j was non-modal in the last q periods, i picks m_ℓ; otherwise, i picks a_ℓ.

I argue that it is a strict sequential equilibrium for all players to follow this strategy, under uniform random matching, for δ and n large enough, and for appropriate q. It will be convenient to define the payoffs received in the sanctions phase $w_1 := g_1(a_1, m_2)$ and $w_2 := g_2(m_1, a_2)$, as well as the corresponding payoffs received by the player carrying out the sanctions $w_1' := g_1(m_1, a_2)$ and $w_2' := g_2(a_1, m_2)$.

First, consider a player i of type 1 who was always modal in the past q periods, and who is only matched with players who were also always modal. If i does not deviate, and n is large enough, her continuation payoff is approximately v_1 for appropriate beliefs about the number of other agents who deviated. (Recall that repeated-game payoffs are normalised by the number of matches s. The player gets v_1 in each of her stage games.) If she deviates she will be non-modal and get at most

$$(1 - \delta)\bar{v}_1 + \delta(1 - \delta^q)w_1 + \delta^{q+1}v_1.$$

So

$$v_1 > (1 - \delta)\bar{v}_1 + \delta(1 - \delta^q)w_1 + \delta^{q+1}v_1 \tag{1}$$

implies that i has no incentive to deviate (for n large enough).

Second, consider a player i of type 1 who was non-modal in the past q periods. (It doesn't matter who she is matched with.) She has the greatest incentive to deviate when there are q periods remaining in the sanctions phase, so suppose that is the case. If i does not deviate, she gets

$$(1 - \delta^q)w_1 + \delta^q v_1.$$

If she deviates, she gets

$$(1 - \delta)\underline{v}_1 + \delta(1 - \delta^q)w_1 + \delta^{q+1}v_1.$$

[12] Note that this is not possible in the standard two-player model, because a player sanctioning another player could get payoffs lower than if *she* were being sanctioned. In that case, it may be optimal for her to deviate to avoid carrying out sanctions. In the present paper, this problem does not arise because of uniform random matching: in any given subgame, players will not expect to be matched with non-modal partners in future periods.

So i has no incentive to deviate provided

$$(1-\delta^q)w_1 + \delta^q v_1 > (1-\delta)\underline{v}_1 + \delta(1-\delta^q)w_1 + \delta^{q+1}v_1,$$

which simplifies to

$$(1-\delta^q)w_1 + \delta^q v_1 > \underline{v}_1. \tag{2}$$

Third, consider a player i of type 1 who was always modal in the past q periods, matched with $r \leq s$ players who were non-modal and $s-r$ players who were always modal. It is sufficient to consider the case where $r = s$, since the case where $r = 0$ is already addressed and payoffs in the other cases are linear combinations of the payoffs in the cases $r = s$ and $r = 0$. Therefore, suppose i is only matched with players who were non-modal. If i does not deviate, she gets w_1' at t. Thereafter, if n is large enough she will expect to get a payoff arbitrarily close to v_1 in every stage game. So for n large enough her expected continuation payoff is arbitrarily close to

$$(1-\delta)w_1' + \delta v_1.$$

If she deviates she gets $\underline{v}_1$ at t, after which she is sanctioned. So her continuation payoff is

$$(1-\delta)\underline{v}_1 + \delta(1-\delta^q)w_1 + \delta^{q+1}v_1.$$

So for n large enough,

$$(1-\delta)w_1' + \delta v_1 > (1-\delta)\underline{v}_1 + \delta(1-\delta^q)w_1 + \delta^{q+1}v_1 \tag{3}$$

implies that she has no incentive to deviate.

Define $v_1^* := (1-\delta^q)w_1 + \delta^q v_1$. We can rewrite the equations as

$$v_1 > (1-\delta)\bar{v}_1 + \delta v_1^*, \tag{4}$$

$$v_1^* > \underline{v}_1, \text{ and} \tag{5}$$

$$v_1 > \frac{1-\delta}{\delta}(\underline{v}_1 - w_1') + v_1^*. \tag{6}$$

Inequalities (4) and (5) are, *mutatis mutandis*, equivalent to inequalities (5) and (7) in Fudenberg and Maskin 1986. Inequality (6) ensures that a player matched with non-modal players has no incentive to deviate. The same argument used by Fudenberg and Maskin can be used to show that there exist δ and q such that inequalities (4) to (6) hold. Therefore it is possible to choose δ, p, and n large enough such that there exists a sequential equilibrium that supports the payoff pair v. □

How much information is necessary to sustain the above equilibrium? I argue that for relevant applications, the amount of information each player requires may not be very large at all. Note that players consider at most $2ps(s+1)$ separate pieces of information in each period.[13] There is no restriction on s, other than it be weakly greater than 3 and a factor of n, so the key determinant of the amount of information required is the players' memory p.

Consider the prisoner's dilemma of figure 1, and suppose we want to construct an equilibrium that supports the cooperative payoff profile $(1, 1)$. What is the minimum p required to sustain such an equilibrium? We can set $a_i = C$ and $m_i = D$ (for $i = 1, 2$ since the game is symmetric). The minmax payoff is $\underline{v}_i = 0$ and the maximum payoff is $\bar{v}_i = 1 + g$. In the sanctions phase, the sanctioned player picks the equilibrium-path action C, while the player carrying out the sanctions picks the minmax action D, so their payoffs are $w_i = -\ell$ and $w'_i = 1 + g$, respectively. Inequality (6) simplifies to

$$-\frac{1-\delta}{\delta}(1+g) - (1-\delta^q)(1+\ell) < 0,$$

which always holds. For $q = 1$, inequalities (4) and (5) simplify to

$$\delta > \frac{g}{1+\ell}, \text{ and}$$
$$\delta > \frac{\ell}{1+\ell}.$$

Therefore in any prisoner's dilemma, deviations only need to be sanctioned for a single period to sustain cooperation for δ large enough. Since we only require $p \geq q$, the equilibrium in the repeated matching game with multiple matching can be supported for $p = 1$. In other words, cooperation can be sustained without players having to remember any further than the previous period.

Intuitively, inequality (6) ensures that players matched with non-modal players have no incentive to deviate. But since in the prisoner's dilemma

[13] There are s actions in each $\boldsymbol{\alpha}_i^t$, and each period players consider p such s-tuples for $s+1$ agents, namely themselves and the players they are currently matched with. In addition, for each action considered, the players also consider the corresponding modal action. So the total number of actions considered, including redundancies, is $2ps(s+1)$. If one considers that players need to know the individual actions that determine the modal actions, then the total is $ps^2(s+1)$. (For each of their and their partners' actions in a given period, players need to know s actions – namely the action and the $s-1$ corresponding actions. So the total is $s(s+1)$ times s times p.)

these players receive the maximum payoff $\bar{v}_i = 1 + g$, they clearly have no incentive to deviate and so the inequality holds for any value of δ and q. Inequality (5) ensures that players who were non-modal have no incentive to deviate. If $q = 1$, these players can gain ℓ by deviating to D but lose $1 + \ell$ in the following period when they are sanctioned, so $\delta > \ell/(1 + \ell)$ implies they will not deviate. Inequality (4) ensures players cannot benefit from deviating on the equilibrium path. When $q = 1$, these players can gain g by deviating and will also lose $1 + \ell$ in the following period, so $\delta > g/(1 + \ell)$ implies they will not deviate.

5. Robustness to noise

I argue that the equilibrium of proposition 1 is robust to noise. Let σ_i be the strategy defined in the proof of proposition 1 for player i, and let σ be the corresponding strategy profile. Let v be the payoff pair sustained by σ. Suppose parameters are such that σ is an equilibrium. Claim 1 is concerned with global stability, as defined by Kandori. Claims 2 to 4 are concerned with persistent noise, in the spirit of Ellison.

An equilibrium is globally stable if from any history, players' continuation payoffs converge to their equilibrium-path values.

Definition 2 (Kandori, 1992)**.** An equilibrium sustaining payoffs $v \in V$ is *globally stable* if for any given finite history of actions h,

$$\lim_{t\to\infty} E(v_i(t)|h) = v_\ell \quad \text{for all } i \in N,$$

where $v_i(t)$ is player i's continuation payoff at t and ℓ is her type.

Global stability implies that the effect of any individual deviation on players' continuation payoffs will eventually die out. It is straightforward to see that the equilibrium σ is globally stable, since punishment phases are finite.

Claim 1. The equilibrium σ is globally stable.

Proof. Consider some history h in period t. From h, players return to the equilibrium-path actions after at most q periods. Therefore for any player i of type ℓ and time $t' > t + q$, $v_i(t') = v_\ell$. So the equilibrium is globally stable. □

Ellison notes that global stability does not tell us anything about the equilibrium under persistent noise. Does it remain an equilibrium? How far are players' expected payoffs from their equilibrium values? I start by introducing two types of persistent noise. First, suppose in each stage game each player *trembles* independently with probability ε_τ. If a player trembles, she plays an action drawn from A at random with uniform probability. Second, suppose there is a fraction ε_ν of noisy players. Noisy players always tremble: in every stage game they play an action drawn from A at random with uniform probability.

I show that the equilibrium σ is robust in this setting. Specifically, I show that σ remains an equilibrium for small enough noise and that the effect of noise on payoffs decreases at least linearly with ε_τ and ε_ν.[14,15]

Claim 2. In the model with trembles and noisy players, σ is a strict, sequential equilibrium, provided ε_τ and ε_ν are small enough. Moreover, let $v(\varepsilon_\tau, \varepsilon_\nu)$ be the expected equilibrium-path payoffs in the model with noise. Then $v - v(\varepsilon_\tau, \varepsilon_\nu) = O(\varepsilon_\tau) + O(\varepsilon_\nu)$ as $\varepsilon_\tau, \varepsilon_\nu \to 0$.

Proof. When ε_τ and ε_ν are small, the players' payoffs following all histories are approximately equal to the payoffs in the case without noise. (Note that this relies on uniform random matching to ensure the noisy players do not disproportionately affect some players.) Since σ is a strict sequential equilibrium in the case without noise, it must be a strict sequential equilibrium in the case with noise, for ε_τ and ε_ν small enough.

Next, consider some player i. For the sake of convenience, suppose i is of type 1. So v_1 is i's equilibrium-path payoff without noise and $v_1(\varepsilon_\tau, \varepsilon_\nu)$ is the equivalent expected payoff with noise. The largest possible effect a single tremble in period t can have on player i's continuation payoff is if it is a tremble by player i herself. In which case the impact is

$$\Delta_\tau \coloneqq (1-\delta^{q+1})v_1 - (1-\delta)\underline{v}_1 - \delta(1-\delta^q)w_1.$$

So the impact of any tremble by any player is at most Δ_τ.

Being matched with a noisy player reduces stage-game payoffs by at most $(v_1 - \underline{v}_1)$, so the impact on average continuation payoffs in each period and

[14] Ellison (1994) considers trembles only. In his case, the proof is more involved because the initial equilibrium is not strict.

[15] Note that the result relies on the assumption that a player cannot observe who her partners' partners are (or at least cannot observe the past actions of her partners' partners). For instance, suppose a player realises that her partner is matched with noisy players and her. Then it may not be a best response to follow the equilibrium strategy. One might conjecture that σ remains an equilibrium provided trembles are the main source of noise, since players would then assign low probability to any particular player being noisy.

each subgame is at most

$$\Delta_\nu \coloneqq (1-\delta)\frac{v_1 - \underline{v}_1}{s}.$$

So i's expected payoff in period 1 is at least

$$\begin{aligned} v_1(\varepsilon_\tau, \varepsilon\nu) &\geq v_1 - \sum_{t=1}^{\infty} 2n\delta^{t-1}\Delta_\tau\varepsilon_\tau - \sum_{t=1}^{\infty} s\delta^{t-1}\Delta_\nu\varepsilon_\nu \\ &= v_1 - \frac{2n\Delta_\tau}{1-\delta}\varepsilon_\tau - \frac{s\Delta_\nu}{1-\delta}\varepsilon_\nu. \end{aligned}$$

Therefore $v_1 - v_1(\varepsilon_\tau, \varepsilon_\nu) = O(\varepsilon_\tau) + O(\varepsilon_\nu)$. □

Next, suppose there are errors in observation. Specifically, for every action or modal action a player observes, with probability ε_ω she observes an action drawn from A at random with uniform probability. Suppose observation errors occur independently. (Note that this is an imperfect private monitoring setting.) Consider for now the case where there are no other sources of noise. Then I argue that σ is no longer an equilibrium for any parameter values. Intuitively, this is because players will attribute any non-modal action by a partner to an observation error rather than a deviation, and will assume other players will not sanction the partner.

Claim 3. In the model with observation errors, σ is not a sequential equilibrium, for any value of ε_ω.

Proof. I focus on the case where ε_ω is vanishingly small, since we are interested in a setting where noise is persistent but rare. Consider a subgame where player i observes a non-modal action by a partner j in the past q periods. According to σ_i, player i should sanction j. However in any sequential equilibrium, i must assign probability one to an observation error rather than to j having deviated. Moreover, since errors are independent she must suppose that no one else will sanction j, and that if she were to sanction j she would be non-modal. Therefore it is not a best response for i to sanction j. So σ is not a sequential equilibrium. □

Claim 3 establishes a strong negative result: the proposed equilibrium may be vulnerable to the presence of observation errors. However in claim 4 I show that when other sources of noise are present, the equilibrium becomes robust to observation errors, provided observation errors are rare relative to the other types of noise: since observation errors are relatively unlikely, players will tend to believe that a non-modal observation is due to a tremble or to a noisy player rather than an observation error, which breaks the argument of claim 3.

Claim 4. In the model with trembles, noisy players, and observation errors, σ is a strict, sequential equilibrium, provided ε_τ and ε_ν are small enough and ε_ω is small enough relative to ε_τ and ε_ν.

Proof. Consider a player i who observes a non-modal action by a partner j as before. If ε_ω is small relative to ε_τ and ε_ν, she will assign most probability to a tremble by j or to j being noisy rather than an observation error. Moreover, since all types of noise are rare, other players will likely sanction j. Therefore player i's payoffs for sanctioning and not sanctioning j are approximately equal to the payoffs in the case without noise. So σ must be an equilibrium. □

The result above relies on the probability of an observation error being vanishingly small relative to other sources of noise, which may not be plausible. However the result does demonstrate that observation errors can be accommodated in at least some circumstances. Moreover, there are other factors that may further mitigate things. In particular, if observation errors are sufficiently correlated between players, the argument that other players won't sanction an observed non-modal action breaks down. In practice, it may be sufficient for observation errors to be somewhat correlated and somewhat rarer than other forms of noise in combination.

6. Conclusion

The analysis above demonstrates how cooperation can be sustained in random matching games without assuming an exogenous labelling mechanism. The result relies on the use of *modal actions*. When players are matched with multiple partners each period, deviations in a given period can be identified with information from that period alone. This significantly reduces the informational burden on players. Moreover, the results are robust to noise and involve strict equilibria.

It is worth noting that the idea of using modal actions to detect deviations can be applied to settings other than random matching games. Consider the following model of intergenerational transfers due to Bhaskar (1998). One player is born each period, and each player lives for two periods. In the first period of a player's life, when she is 'young', she receives an endowment. The 'old' player in a given period does not have an endowment, but she can receive transfers from the young player of the following generation. (As in an unfunded pension system.) The socially optimal outcome is for young players to transfer half of their endowment, so that consumption is smoothed over time. Samuelson (1958) noted that this outcome could be supported by a Nash equilibrium. Bhaskar shows that it cannot be supported by a subgame-perfect equilibrium if players have limited memory: players who deviate must be sanctioned by the following generation but cannot be distinguished from players who sanction a deviator; as a result, any candidate equilibrium unravels by backward induction. However, if multiple people are born each period one can determine whether a player deviated by comparing her actions to those of other players born in the same period, so limited memory is not a problem. Conceptually, random matching games and models of overlapping generations are similar. In both cases, players who observe a deviation directly are unable to carry out sanctions, so the problem is to have sanctions carried out by third parties. Modal actions provide a solution to this problem by allowing effective information transmission about which players to sanction.

Modal actions can also be applied to standard repeated games with bounded memory. As mentioned in section 1, existing bounded-memory folk theorems tend to involve complex strategies. The key problem is to construct strategies such that deviations can be detected and the identity of the deviator revealed. Barlo et al. (2009) call such strategies 'confusion-proof'. In Barlo et al. 2016, players play specific sequences of actions, which the authors call 'signalling sequences', at regular intervals. These serve as reminders of the play path and tell players who should be sanctioned. Players need to be incentivised to play the signalling sequences, which renders the strategies complex. Moreover, the sequences need to be spaced far apart to ensure that they are not too costly, so that the required memory is potentially large. In contrast, modal actions can be used to derive a simple folk theorem in pure strategies, provided the stage game is symmetric. If the game is symmetric and players strategies are symmetric, then any deviating player will be non-modal, so deviations can be easily detected even with a bounded memory. In this example, as in the overlapping generations model and in the case of repeated games with random matching, the key conceptual point is that modal actions allow deviations to be succinctly encoded using information about actions from the period in which they occurred.

References

Robert J. Aumann and L. S. Shapley. Long-term competition: A game-theoretic analysis. Working paper, 1976.

Robert Axelrod and William D. Hamilton. The evolution of cooperation. *Science*, 211(4489):1390–1396, 1981.

Mehmet Barlo, Guilherme Carmona, and Hamid Sabourian. Repeated games with one-memory. *Journal of Economic Theory*, 144(1):312–336, 2009.

Mehmet Barlo, Guilherme Carmona, and Hamid Sabourian. Bounded memory folk theorem. *Journal of Economic Theory*, 163:728–774, 2016.

Venkataraman Bhaskar. Informational constraints and the overlapping generations model: Folk and anti-folk theorems. *Review of Economic Studies*, 65(1):135–149, 1998.

Joyee Deb. Cooperation and community responsibility. Working paper, 2019.

Joyee Deb and Julio González-Díaz. Enforcing social norms: Trust-building and community enforcement. Working paper, 2019.

Francesc Dilmé. Helping behavior in large societies. *International Economic Review*, 57(4):1261–1278, 2016.

Glenn Ellison. Cooperation in the prisoner's dilemma with anonymous random matching. *Review of Economic Studies*, 61(3):567–588, 1994.

Drew Fudenberg and Eric Maskin. The folk theorem in repeated games with discounting or with incomplete information. *Econometrica*, 54:533–554, 1986.

Drew Fudenberg and Eric Maskin. On the dispensability of public randomization in discounted repeated games. *Journal of Economic Theory*, 53(2): 428–438, 1991.

Joseph E Harrington. Cooperation in a one-shot prisoners' dilemma. *Games and Economic Behavior*, 8(2):364–377, 1995.

Yuval Heller and Erik Mohlin. Observations on cooperation. *Review of Economic Studies*, 85(4):2253–2282, 2017.

Michihiro Kandori. Social norms and community enforcement. *Review of Economic Studies*, 59:63–80, 1992.

Masahiro Okuno-Fujiwara and Andrew Postlewaite. Social norms and random matching games. *Games and Economic Behavior*, 9(1):79–109, 1995.

Ariel Rubinstein. Equilibrium in supergames with the overtaking criterion. *Journal of Economic Theory*, 21:1–9, 1979.

Paul A Samuelson. An exact consumption-loan model of interest with or without the social contrivance of money. *Journal of Political Economy*, 66: 467–482, 1958.

William H Sandholm. *Population games and evolutionary dynamics*. MIT Press, Cambridge, MA, 2010.

Robert Sugden. *The Economics of rights, co-operation and welfare*. Basil Blackwell, Oxford, 1986.

Satoru Takahashi. Community enforcement when players observe partners' past play. *Journal of Economic Theory*, 145(1):42–62, 2010.

Jörgen W Weibull. *Evolutionary game theory*. MIT Press, Cambridge, MA, 1995.