Belief free equilibria[1]

Olivier Compte and Andrew Postlewaite

First version: January 2013. This version: April 2013

Abtract: The repeated game literature studies long run/repeated interactions, aiming to understand how repetition may foster cooperation. Conditioning future behavior on past play is crucial in this endeavor. For most situations of interest a given player does not directly observe the actions chosen by other players and must rely on noisy signals he receives about those actions. This is typically incorporated into models by assuming that there is a monitoring structure that consists of a joint probability distribution over the signals each player receives, for each given actions players may choose. Although this is meant simply to capture the fact that players don't directly observe the actions chosen by others, constructed equilibria often depend on players precisely knowing the distributions, somewhat unrealistic in most problems of interest. This paper aims to show the fragility of belief free equilibrium constructions when one adds shocks to the monitoring structure in repeated games.

1 Introduction

The repeated game literature studies long run/repeated interactions, aiming to understand how repetition may foster cooperation. Conditioning behavior on observations is an essential ingredient, and the literature has tried to understand how the nature and quality of observations affect cooperation possibilities. Specifically, the analysis starts with a stage game characterized by a payoff structure describing how action profiles affect gains for each player, and a monitoring structure, that is, a joint distribution over signal and action profiles, capturing the possibility that the actions played are typically not perfectly observable. Taking payoff and monitoring structures as given, one then attempts to characterize the set of (sequential) equilibria of the repeated game.

[1] We thank Stephen Morris, Larry Samuelson and Yuichi Yamamoto for helpful comments.

Although an imperfect monitoring structure is just a modelling device employed to capture an agent's inability to observe perfectly what others are doing, equilibrium constructions sometimes hinge on the exact specification of that monitoring structure, with strategies finely tuned to that particular specification, as though agents could easily determine the precise (stochastic) relationship between the action profile played and the signals observed, despite the fact that others' actions are not observable and that others' signals may not be observable either. This seems unrealistic, more so when signals are privately observed. But even if one accepts that assumption, one may question the robustness of equilibrium constructions obtained in this way. One suspects that the plethora of equilibria one can construct might be a consequence of this presumed unlimited ability of agents to tailor their strategies to the underlying parameters of the game.

This paper illustrates the lack of robustness of some of these equilibrium constructions, by considering an environment in which there are exogenous and persistent shocks to the monitoring structure. These shocks prevent players from tailoring equilibrium strategies to one particular monitoring structure. We consider the belief free construction proposed by Ely and Valimaki (2002) (EV hereafter) to support cooperation in the prisoner's dilemma game, and show that except for isolated parameters, the EV construction does not allow one to support cooperation in the environment that we propose. We next show that our argument extends to the more elaborate belief free construction proposed by Piccione (2002), in which strategies are allowed to depend on longer and possibly infinite histories.

2 The model

We consider a standard repeated prisoner's dilemma, with discount equal to δ. At each date, each player chooses one of two actions, Cooperate or Defect: $a_i \in A_i \equiv \{C, D\}$. We denote by g^a the expected gains associated with action profile $a = (a_1, a_2)$. After choices have been made, each player receives a private signal $y_i \in Y_i$. The signal profile $y = (y_1, y_2)$ is assumed to be correlated with the action profile a, thus defining a monitoring structure. In contrast to traditional approaches to repeated games however, the monitoring structure is not assumed to be identical across periods. Rather, there is an underlying state $\theta \in \Theta$ that captures variations in the monitoring structure. Specifically, we denote by $q_a^\theta(y)$ the probability that y is observed when a is played, and when the state is θ. We also assume that θ follows a Markov process, characterized by transition probabilities $\pi_{\theta\theta'}$. Importantly, the state θ is not observed by players, though the signal received may provide information about θ.

Memory 1 strategies. We next define *memory 1 strategies* as strategies in which a player's behavior depends only on last period own action a_i and signal y_i. Formally, let $H_i^1 = A_i \times Y_i$ denote the set of private histories of length 1. A memory 1 strategy is characterized by a vector of probabilities $p_i \equiv \{p_i^{h_i}\}_{h_i \in H_i^1}$, in which $p_i^{h_i}$ refers to the probability that player i cooperates after history h_i.

A key observation is that when both players use memory 1 strategies, continuation payoffs depend only on the current state θ and the current action profile a played.[2] Having fixed a candidate equilibrium p, we denote by $v_i^{\theta,a}$ the continuation values induced by p.

The EV construction. The EV construction then relies on considering equilibria in which for each player, and after any possibly long history of the game, the incentives to play C or D are weak (and thus independent of previous history of play). A priori, this would seem to generate many constraints. However, the observation above plays a key role. When θ is fixed, one only has to check the following incentive conditions:

$$\text{For all } i, j \neq i,\ v_i^{\theta,C,a_j} = v_i^{\theta,D,a_j} \text{ for } a_j = C, D. \tag{1}$$

Indeed, when these incentive conditions hold, player i's belief about the action that player j plays in the current period is possibly relevant to player i's continuation payoff, but it is irrelevant to his incentives: player i remains

[2]This is because a and θ determine the distribution over signal profile $y = (y_1, y_2)$ received, and because next period play for each player i will then depend solely on $h_i = a_i y_i$.

indifferent between cooperating and defecting whatever choice of player j – the equilibrium obtained is belief free.

How restrictive are these conditions? There are 2 incentive constraints for each player. With at least 2 possible signals being received, a strategy is defined by at least 4 parameters. That leaves many degrees of freedom, and many equilibria can thus be constructed in this way.

Varying monitoring structure.

When the monitoring structure varies over time and yet has some persistence, even very old signals may provide information about the current monitoring structure (i.e. about θ), and beliefs about the monitoring structure may affect incentives. Applying the logic of the belief free approach, one may circumvent this difficulty by looking for equilibria in which incentives to play C or D are weak for all possible realizations of θ. In this varying monitoring environment, we shall say that an equilibrium supporting cooperation is belief free if and only if[3]

$$\textit{For all } \theta \text{ and } i, j \neq i,\ v_i^{\theta,C,a_j} = v_i^{\theta,D,a_j} \text{ for } a_j = C, D. \tag{2}$$

It should be obvious however that as the number of distinct states increases, the number of constraints will exceed the number of free parameters, so generically, such an equilibrium cannot exist.

To be more precise, we write down some of the relationships between payoff and monitoring structures imposed by these constraints. Given a strategy p_j played by player j, define $\phi^a_{p_j,\theta}$ as the probability that player j cooperates in the next period when the action profile a is played in the current period and the current state is θ:

$$\phi^a_{p_j,\theta} = \sum_{y=(y_1,y_2)\in Y} q^\theta_a(y) p_j^{a_j y_j}.$$

The incentive conditions allow us to define $v_i^{\theta,a_j} \equiv v_i^{\theta,C,a_j} = v_i^{\theta,D,a_j}$. We also let $\overline{v}_i^{\theta,a_j}$ denote the expected payoff obtained by player i when player j plays

[3] The state-by-state constraints are analogous to the ex post equilibrium constraints introduced in Fudenberg and Yamamoto (2010).

a_j and last period state is θ:

$$\overline{v}_i^{\theta,a_j} = \sum_{\theta'} \pi^{\theta\theta'} v_i^{\theta',a_j}.$$

We have:

$$v_i^{\theta,a_i,a_j} = (1-\delta)g^a + \delta[\phi^a_{p_j,\theta}\overline{v}_i^{\theta,C} + (1-\phi^{a,C}_{p_j,\theta})\overline{v}_i^{\theta,D}]. \tag{3}$$

Computing $v_i^{\theta,C,a_j} - v_i^{\theta,C,a_j}$, the incentive condition may thus be rewritten as:

$$(1-\delta)(g^{Ca_j} - g^{Da_j}) = \delta(\phi^{Da_j}_{p_j,\theta} - \phi^{Ca_j}_{p_j,\theta})(\overline{v}_i^{\theta,C} - \overline{v}_i^{\theta,D}).$$

Setting $\rho = \frac{g^{CC}-g^{DC}}{g^{CD}-g^{DD}}$, one obtains the following linear relationships:

$$\phi^{DC}_{p_j,\theta} - \phi^{CC}_{p_j,\theta} - \rho(\phi^{DD}_{p_j,\theta} - \phi^{CD}_{p_j,\theta}) = 0. \tag{4}$$

When the number of distinct states exceeds $2 \mid Y \mid$, these equations cannot hold simultaneously for all states, at least for generic monitoring structures.[4]

Longer histories.

The argument above extends to the more elaborate constructions proposed in Piccione (2002). In that construction, players condition behavior on longer histories, but conditioning on longer histories does not necessarily imply more degrees of freedom in choosing the probabilities $p_i^{h_i}$.

Formally, we now define H_i as the set of relevant histories for player i, where by relevant we mean that each distinct $h_i \in H_i$ gives rise to a different continuation behavior $\sigma_i^{h_i}$ for player i. For the memory-1 strategies with two signals $y_i \in \{0,1\}$, there are just 4 relevant histories: $H_i = \{C1, C0, D1, D0\}$. In Piccione's construction, h_i is defined as the number

[4]The condition could actually be strengthened, as we have only used a subset of the equations that need to hold in equilibrium.

Note that we have assumed payoffs to be independant of the monitoring structure. If one assumes that each player gets a payoff $r_i(a_i, y_i)$, we would obtain expected payoffs that depend on the monitoring structure, thus implying that we define $g_i^{\theta,a} = E_{\theta,a} r_i(a_i, y_i) = \sum_y q_a^\theta(y) r_i(a_i, y_i)$. We would then obtain polynomial (rather than linear) relationships between the monitoring structure $(q^\theta)_\theta$ and the strategy profile p.

of consecutive bad signals (say $y_i = 0$) received since the last good signal ($y_i = 1$). So $H_i = \mathcal{N}$.

A strategy can then be characterized as before by a vector $p_i \equiv \{p_i^{h_i}\}_{h_i \in H_i}$, where $p_i^{h_i}$ is the probability of cooperating after h_i. Taking p_j as given, we next denote by $v_i^{a_i,h_j}$ the (continuation) value that i obtains when he plays a_i and the current (relevant) history for j is h_j. The belief free conditions require

$$v_i^{C,h_j} = v_i^{D,h_j}.$$

There are as many constraints as unknowns, but all these constraints are not independent. We have seen that in the EV construction, with memory 1 strategies, there are only two independent constraints.[5] In Piccione's construction, the constraints are more stringent and leave only one degree of freedom: once $p_j^{h_j}$ has been chosen for $h_j = 0$, all other $p_j^{h_j}$ are uniquely determined through a sequence of recursive equations.[6] So Piccione's construction fails to be robust as well.

Of course in general, we should expect that conditioning on longer histories would increase the number of degrees of freedom one is left with once incentive conditions have been taken care of. For strategies with bounded memory for example, the number of independent constraints could be as small as (but not smaller than) $\frac{|H_i|}{|Y_i|}$.[7] Still, when the monitoring structure varies, each of these constraints generically gives rise to $\mid \Theta \mid$ constraints, so equilibrium constructions typically fail to be robust when $\mid \Theta \mid > \mid Y_j \mid$.

[5] This is because $v_i^{a_i,h_j} = p_j^{h_j} v_i^{a_i,C} + (1 - p_j^{h_j}) v_i^{a_i,D}$, where $v_i^{a_i,a_j}$ is defined as before.

[6] The recursive equations are:

$$v_i^{a_i,h_j} = (1-\delta) g^{a_i,p_j^{h_j}} + \delta (q_{C,p^j} v_i^{a_i,0} + (1 - q_{C,p^j}) v_i^{a_i,h_j+1})$$

where $q_{C,h^j} = \Pr_{C,p_j^{h_j}}(y^j = 1)$.

Defining $v_i^{h_j} \equiv v_i^{C,h_j} = v_i^{D,h_j}$, p_j^0 determines v_i^0 and v_i^1, which then determines p_j^1, which in turn determines v_i^2, and so on.

[7] This is because for a given bound on memory, $v_i^{a_i,h_j}$ is fully determined by the current action played by j and the truncated history $\check{h}_j$ that excludes the oldest action and signal: $v_i^{a_i,h_j} = p_j^{h_j} v_i^{a_i,(\check{h}_j,C)} + (1 - p_j^{h_j}) v_i^{a_i,(\check{h}_j,D)}$, where $v_i^{a_i,(\check{h}_j,a_j)}$ is precisely the value that i obtains when the truncated history is $\check{h}_j$ and the current action is a_j. The number of distinct pairs $(\check{h}_j, a_j)$ is presicely equal to $\frac{|H_i|}{|Y_i|}$.

Rich set of monitoring structures

We finally consider the case of a *rich set* Θ of monitoring structures. By this we mean the following: we fix a monitoring structure q_0 having full support, and assume that Θ contains a ball B of monitoring structures around q_0. To simplify notation, we identify θ with its associated monitoring structure q^θ. To avoid unnecessary technicalities, we also assume that the monitoring structure does not change over time ($\pi_{\theta\theta} = 1$). We provide below a general argument showing that there cannot exist robust belief free equilibria when the set of monitoring structures is rich (even when there are no bounds imposed on memory).

Consider a candidate belief free equilibrium represented by a vector (p_i, p_j) where $p_j \equiv \{p_j^{h_j}\}_{h_j \in H_j}$. Having fixed p_j, we consider the value $v_i^{a_i,h_j}(q)$ that player i obtains under monitoring structure $q \in B$ when he always plays a_i and his opponent's history is h_j. By definition, in a robust belief equilibrium

$$v_i^{C,h_j}(q) = v_i^{D,h_j}(q) \text{ for all } h_j \text{ and } q \in B.$$

Now consider two signals $\overline{y}_j$ and $\underline{y}_j$, $\Delta > 0$, and a monitoring structure q having the following properties: (i) for all $a \neq CC$, $q_a(y) = q_a^0(y)$; (ii) for $a = CC$, and for all y_i and $y_j \neq \overline{y}_j, \underline{y}_j$, the marginals $q_{CC}(y_k)$ are unchanged, i.e. $q_{CC}(y_k) = q_{CC}^0(y_k)$ (ii) for signals $y_j = \overline{y}_j, \underline{y}_j$, $\mid q_{CC}(y_j) - q_{CC}^0(y_j) \mid = \Delta > 0$; (iii) $q \in B$. This is possible for Δ small enough because q_0 has full support and because the constraints on marginals impose fewer constraints than there are unknowns.

By construction $v_i^{D,h_j}(q) = v_i^{D,h_j}(q^0)$, so the belief free condition implies

$$v_i^{C,h_j}(q) = v_i^{C,h_j}(q^0) \equiv v_i^{C,h_j}.$$

Now observe that this condition holds for histories $(h_j, C\overline{y}_j)$ and $(h_j, C\underline{y}_j)$ as well, so, using the usual recursive equations, we can write:[8]

$$0 = v_i^{C,h_j}(q) - v_i^{C,h_j}(q^0) = \delta(q_{CC}(\overline{y}_j) - q_{CC}^0(\overline{y}_j))p_j^{h_j}(v_i^{C,(h_j,C\overline{y}_j)} - v_i^{C,(h_j,C\underline{y}_j)})$$

[8]Note that this equality holds even if the gain g_a depends on the monitoring structure through the signal y_i received by player i, because the marginals on y_i are kept unchanged.

implying that whenever $p_j^{h_j} > 0$, $v_i^{C,(h_j,C\overline{y}_j)} = v_i^{C,(h_j,C\underline{y}_j)}$. Applying the same argument to a monitoring structure q' that only differs from q_0 on $q_{CD}(\overline{y}_j)$ and $q_{CD}(\underline{y}_j)$, one obtains that if $1 - p_j^{h_j} > 0$, then $v_i^{C,(h_j,D\overline{y}_j)} = v_i^{C,(h_j,D\underline{y}_j)}$. We thus conclude that player i's continuation payoff does not depend on the signal received by player j, so player i has strict incentives to defect.

3 Discussion

Although we have focused on belief free equilibria because the point is particularly easy to make in that case, the general methodological point that we raise applies to the more standard equilibrium constructions, including those obtained in games with imperfect public monitoring (Fudenberg, Levine, and Maskin (1994)). We pursue this path in a companion paper (Compte and Postlewaite (2013)). Adopting the framework above with a varying monitoring technology (again characterized by an unobservable underlying state θ), and asking that incentives hold for all possible realizations of that state θ, one obtains tighter constraints that are shown to generate inefficiencies when monitoring is imperfect and the set of states is rich enough, even as players become arbitrarily patient. In that application, it is not the equilibrium construction itself that fails to be robust. Rather, it is the possibility of sustaining a given Pareto efficient point that fails to be robust: because equilibrium strategies cannot be tailored to each different realization of the state, the locus of each state contingent equilibrium values drifts away from the point that one would have liked to sustain.

Finally we mention the work of Bhaskar (2000) and Bhaskar, Mailath and Morris (2008) who also aim to show the lack of robustness of belief free equilibria. Following Harsanyi (1973), the path taken by these papers consists of asking whether equilibria can be purified with payoff shocks. They show that there is a failure of robustness with finite memory strategies for public monitoring games, but do not obtain conclusive results for private monitoring games. One difference between their approach and ours is that purification is essentially a weaker robustness requirement. Purification arguments ask whether one can find a *nearby* equilibrium for a *nearby* perturbation, hence implicitly assuming that the variation in the common prior (i.e. the perturbation) is observable. In contrast, when one looks at a varying environment as we do, with variations not directly observable, and when one asks for belief free equilibria despite these variations, one is actually asking for a *given* strategy to remain an equilibrium even as the monitoring structure changes.

References

Bhaskar, V. (2000), "The Robustness of Repeated Game Equilibria to Incomplete Payoff Information," mimeo, University of Essex.

Bhaskar, V., G. Mailath, and S. Morris (2008), "Purification in the Infinitely-Repeated Prisoners' Dilemma", *Review of Economic Dynamics*, 11, 515-528.

Compte, O., and A. Postlewaite (2013), "Folk Theorems", mimeo.

Ely, J. C., and J. Välimäki (2002), "A Robust Folk Theorem for the Prisoner's Dilemma," *Journal of Economic Theory*, 102, 84-105.

Fudenberg, D., D. K. Levine, and E. Maskin (1994), "The Folk Theorem with Imperfect Public Information," *Econometrica*, 62, 997-1040.

Fudenberg, D., and Y. Yamamoto (2010), "Repeated Games Where the Payoffs and Monitoring Structure Are Unknown," *Econometrica*, 78, 1673-1710.

Harsanyi, J. C., (1973), "Games with Randomly Disturbed Payoffs: A New Rationale for Mixed-Strategy Equilibrium Points," *International Journal of Game Theory* 2, 1–23.

Piccione, M. (2002), "The Repeated Prisoner's Dilemma with Imperfect Private Monitoring," *Journal of Economic Theory*, 102, 70-83.