Codes of Conduct, Private Information and Repeated Games*

Juan I. Block† David K. Levine‡

August 17, 2015

Abstract

We examine self-referential games in which there is a chance of understanding an opponent's intentions. Even when this source of information is weak, we are able to prove a folk-like theorem for repeated self-referential games with private monitoring. Our main focus is on the interaction of two sources of information about opponents' play: direct observation of the opponent's intentions, and indirect observation of the opponent's play in a repeated setting.

Keywords: Repeated game, folk theorem, self-referential game, approximate equilibrium
JEL classification: D01 D03 D82

*We would like to thank Drew Fudenberg, Phillip Johnson, George Mailath, Wolfgang Pesendorfer, Balazs Szentes, and Satoru Takahashi for helpful comments. Juan Block acknowledges support from the Cambridge-INET Institute, and David Levine thanks the National Science Foundation (Grant SES-0851315) and the European University Institute for financial support.

†Faculty of Economics, University of Cambridge. Email: jb2002@cam.ac.uk

‡Departments of Economics, European University Institute and Washington University in St. Louis. Email: david@dklevine.com

1 Introduction

The theory of repeated games has made enormous strides in penetrating the difficult but relevant setting in which players observe noisy signals of each other's play.[1] Unfortunately as our knowledge of equilibria in these games has expanded there is an increasing sense that the types of equilibria studied – involving as they do elaborately calibrated indifference – are difficult for players to play and unlikely to be observed in practice. By way of contrast, the theory of approximate equilibria in repeated games is simpler and generally more satisfactory than the theory of exact equilibrium.[2] Moreover, the abstract world of repeated games is not very like the world we inhabit. In poker games players can never guess that their opponent is bluffing from the expression on his face, and skilled interrogators who by asking a few pointed questions can never tell whether a suspect is lying or telling the truth.

In this paper we examine economic situations where the *hypotheses of involuntary truth-telling*, that is the ability to detect intentions (Schelling (1978), and Gauthier (1986)), and the use of intentions as conditional commitment devices (Frank (1988), and Hirshleifer (1987)) play a significant role on determining agents' behavior. There exists a great many experimental studies both in economics and psychology showing a positive effect of intention recognition (see, for instance, Zuckerman et al. (1981) and Sally (1995) for a survey on this subject). The challenge for modeling this idea is to capture both the voluntary report of intentions and the lack of self-control about hiding intentions (that is, the implicit cost of lying).

Our setup includes a class of games in which players have at least a chance of fathoming each other's intentions as in Levine and Pesendorfer (2007) but allowing for asymmetries and more than two players. It utilizes the notion that players employ codes of conduct which are defined as a complete specification of how they play and their opponents "should" play. This in turn allows us to have a well-defined agreement on social norms when there are multiple agents with different strategies. Players also receive signals about what code of conduct their opponent may be using, while their own code of conduct enables them to respond to these signals. This is the *self-referential* nature of the games studied here.

An effective code of conduct rewards players for using the same code of conduct, and punishes them for using a different code of conduct. Several examples explore such issues as when players in a repeated setting might get information about the past play of new partners from other players. General results about when approximate equilibria in a base game can be sustained as strict equilibria in the corresponding self-referential game are given. As an application a discounted strict Nash folk-like theorem for enforceable mutually punishable payoffs in repeated games with private monitoring is proven despite very limited ability to observe directly codes of conduct. Our basic conclusion is that direct observation of opponents' intentions and repetition of a game are proven to enhance players' understanding of individuals' past behavior in noisy environments. Even if direct observation is unreliable, it may be enough to overcome the small ε's that arise when simple repeated game strategies are employed.

[1] See, for example, Lehrer (1990), Compte (1998), Kandori and Matsushima (1998), Matsushima (2004), Ely et al. (2005), Hörner and Olszewski (2006), Fudenberg and Levine (2007), and Sugaya (2011).

[2] It is difficult to rationalize, for example, why a player who is aware that opponent has very favorable signals about his behavior, does not take advantage of this knowledge to behave badly. It is exactly this type of small gain that approximate equilibrium constructions are based on (see Fudenberg and Levine (1991), and Renou and Tomala (2013)).

1.1 Related Literature

Our paper relates to the literature on conditional commitment devices, Tennenholtz (2004), Kalai et al. (2010), and Peters and Szentes (2012). The construction of "program equilibria" (Tennenholtz (2004)) requires players submitting computer programs that takes as input the opponent's computer program, thereby implementing any outcome where players receive at least their minmax payoffs. Unlike our results, this folk theorem holds only for pure strategies and in fact computer algorithm are a special case of the code of conduct space considered here. A more precise characterization of mutually dependent commitment devices was proposed by Kalai et al. (2010), who by incorporating mixed strategies and a richer set than Turing machines show a folk theorem sustaining any correlated outcome above the minmax payoffs. In Bayesian Games, Peters and Szentes (2012) examine these commitment devices (formally modeled as definable functions) and show that equilibrium payoffs cannot be worse than agents' minmax point. Our work is also related to the literature on common agency problems that applies self-referential contracts such as Szentes (2015) and Peters (2015). Our approach differs from this literature in a fundamental aspect – we focus on noisy environment by allowing agents to observe imperfectly informative signals about each other conditional commitment devices.

Most closely related is Bachi et al. (2014) who study two player games in which deceptive agents may betray their true intentions. They find that if the cost of deception is sufficiently high the payoff set may expand. Although they explore the role of signals about potential behavior, their analysis only considers agents choosing a deceptive strategy which is observed with some probability, and leaves aside the case in which agents intend to play truthfully but their opponent observes a signal that is associated with lying. We study games with more than two players and are able to characterize the crucial issue of communication with non-biding messages between detectors and punishers. Levine and Pesendorfer (2007) examined self-referential games as a simple alternative to repeated games that exhibit many of the same features. Their finding is that in a two player symmetric setting if players can accurately determine whether or not their opponent is using the same strategy as they are then a type of folk theorem holds. But, they focus on a simple class of games in order to identify which of many equilibria have long-run stability properties in an evolutionary setting.

The paper proceeds as follows. Section 2 presents the model. In Section 3 we illustrate self-referential games with various examples. In Section 4 we provide our main theoretical application of self-referential games. We conclude in Section 5. The Appendix collects an auxiliary result.

2 The Model

2.1 The Base Game

Consider a finite *base game* with set of players $I = \{1, \dots, N\}$. Each player i chooses a strategy s_i from the strategy set S_i.[3] Let $s \in S := \times_i S_i$ be the profile of strategies. The preferences of player i are represented by a von Neumann-Morgenstern utility function $u_i : S \to \mathbb{R}$ with the usual notation $u_i(s_i, s_{-i})$ where $s_{-i} \in S_{-i} := \times_{j \neq i} S_j$ is used for the strategy profile of all players but player i. We denote by $\Gamma = \{I, (S_i, u_i)_{i \in I}\}$ this base game. For any $\varepsilon \geq 0$, $s \in S$ is an ε-Nash equilibrium if for all i and $\tilde{s}_i \neq s_i$ it holds that $u_i(s_i, s_{-i}) \geq u_i(\tilde{s}_i, s_{-i}) - \varepsilon$.

[3] We allow mixed strategies, but there are only a finite number of them. In repeated games, we assume implicitly either a finite horizon, or a very small subset of strategies in an infinite horizon. For instance, finite automata with an upper bound on the number of states.

2.2 The Self-Referential Game

For any base game Γ, we can define the *self-referential game* $G(\Gamma)$. Every player i privately observes a signal $y_i \in Y_i$, where Y_i is finite, and $y \in Y := \times_i Y_i$ is the set of private signal profiles. The strategy of player i in $G(\Gamma)$ is referred to as a *code of conduct,* denoted by r^i. Each player i is endowed with a common space of codes of conduct $R_0 = \{r^i \in \mathbb{R}^N | r^i_j : Y_j \to S_j \text{ for all } j\}$, where each r^i_j defines a mapping from Y_j to S_j.[4] We write $r \in R := \times_i R_0$ for the profile of codes of conduct. Essentially, each player chooses a code of conduct that everyone is supposed to follow (what to play conditional on each private signal y_j) and simultaneously commits to follow the code himself.

For each profile of codes of conduct $r \in R$, let $\pi(\cdot|r)$ be the probability distribution over signal profiles Y. The collection of probability distributions over profile of private signals is denoted by $\{\pi(\cdot|r)|r \in R\}$. Let $\pi_i(\cdot|r)$ be the marginal probability distribution of $\pi(\cdot|r)$ over Y_i. That is, $\pi_i(y_i|r)$ is the probability that player i observes $y_i \in Y_i$ if players have chosen profile of codes of conduct $r \in R$.

Notice that codes of conduct play two fundamental roles. First, they determine how players behave as a function of the signals they receive (the self-reports about intentions), that is, a player who has chosen the code of conduct r^i and who observes the signal y_i plays according to $r^i_i(y_i) = s_i$. Second, codes of conduct influence the signals y players receive about each others' intentions through the probability distribution π (an involuntary signal about behavior).

If the profile of codes of conduct is $r \in R$ the expected utility of player i is given by

$$U_i(r) = \sum_{y \in Y} u_i(r^1_1(y_1), \ldots, r^N_N(y_N))\pi(y|r).$$

A Nash equilibrium of the self-referential game is a profile of codes of conduct $r \in R$ such that for all players i and any alternative code of conduct $\tilde{r}^i \neq r^i$, $U_i(r^i, r^{-i}) \geq U_i(\tilde{r}^i, r^{-i})$.

The timing of G is as follows. Before playing Γ and observing any Y_i, players simultaneously choose r^i and are committed to follow it. Afterwards, each player i privately observes y_i generated by $\pi(y|r)$ for $y \in Y$ given $r \in R$. Finally, players execute $r^i_i(y_i) = s_i$ for $s_i \in S_i$.

3 Examples

The main purpose of the following examples is to illustrate how self-referential codes of conduct work. Let the self-referential game G consist of a binary space of signals $Y_i = \{0, 1\}$ for each player i, and a space of codes of conduct $R_0 = \{r^i | r^i_j : \{0, 1\} \to S_j \text{ for all } j = 1, 2\}$.[5] The marginal probability distribution of signals is:

$$\pi_i(y_i = 1|r) = \begin{cases} q & \text{for } r^1 \neq r^2, \\ p & \text{otherwise.} \end{cases}$$

Where $\pi(y|r) = \pi_i(y_i|r)\pi_j(y_j|r)$ for all players i, j and $q \geq p$.[6] In words, signal 0 may be interpreted as "we are both using the same code of conduct" and signal 1 may be interpreted as "we are both

[4]We assume a common space of codes of conduct, but we can allow for heterogeneous strategy spaces. In that case several strategies might induce the same map while deferring in terms of the probability distribution of signals.

[5]Private signals could represent verbal and nonverbal communication between the agents, for example, simple cues as handshake, winks and smiles, or voluntary promises (Charness and Dufwenberg (2006)).

[6]That is, private signals are conditionally independent thereby implying that the opponent j's private signal has no information about whether player i observes y_i when players choose codes of conduct.

Table 1: Prisoner's dilemma payoff matrix

	C	D
C	5, 5	0, 6
D	6, 0	1, 1

using different codes of conduct." Importantly, the probability distribution of signals is defined for *asymmetric* games but players can recognize whether their opponent plays the *same* code of conduct or not. We maintain G throughout the next two examples.

3.1 Prisoner's Dilemma

We will study a prisoner's dilemma game. The actions are denoted C for cooperate and D for defect, and the payoffs are given in Table 1. We focus on pure strategies $S_i = \{C, D\}$ for all players i and we work with it as the base game Γ.

One equilibrium profile of codes of conduct r is simply to ignore the signal and defect, that is, all players $i = 1, 2$ choose strategy r^i such that $r^i_j(y_j) = D$ for any signal y_j.[7] This is a Nash equilibrium of the self-referential game G exactly as in the strict Nash equilibrium of the base game Γ, and each player gets 1.

Let us investigate the possibility of sustaining cooperation through self-referentiality, exploiting G. The fact that players can commit to codes of conduct does not necessarily imply that cooperation can be sustained with certainty. If one agent chooses to always cooperate, his opponent may choose an alternative code of conduct whereby always defecting and guaranteeing himself a payoff 6 while the cooperator getting 0. Consequently, we consider the code of conduct $\hat{r}^i$ that chooses the cooperative action C if the signal 0 is received, thereby agents agreeing that there is room for cooperation whenever they expect their opponents to also consider the informative signals, and that chooses D if the signal 1 is received. Formally each player i chooses $\hat{r}^i$ given by

$$\hat{r}^i_j(y_j) = \begin{cases} C & \text{for } y_j = 0, \\ D & \text{otherwise.} \end{cases}$$

If both players adhere to $\hat{r}^i$, they receive an expected utility of $U_i(\hat{r}) = 5 - 4p$. Of course, signals are noisy but the more informative the signal structure of G is when they agree on the code of conduct (the lower p), the closer is their expected payoff to the efficient payoff vector. A player who chooses instead to always defect, thus $\tilde{r}^i_i(y_i) = D$ for all y_i and $\tilde{r}^i_j(y_j) = s_j$ for any s_j, y_j, gets $U_i(\tilde{r}^i, \hat{r}^j) = 6 - 5q$, and does worse by always cooperating as we mentioned above. The code of conduct profile $\hat{r}$ is a Nash equilibrium of G when $q \geq 1/5 + 4/5p$. This says, in effect, that the signal must be informative enough if cooperation is to be sustained.

3.2 Repeated Prisoner's Dilemma

We now consider the prisoner's dilemma repeated twice and we use the sum of payoffs between the two periods. To facilitate the analysis with respect to the one-shot version of the game let us focus again on pure strategies, $S_i = \{C, D\}^2$. Consider the following code of conduct r^i: $r^i_j(y_j = 0) = CC$ and $r^i_j(y_j = 1) = DD$. Since play is not conditioned on what the other player does in the first period, the optimal alternative code of conduct $\tilde{r}^i$ against this code r^i is given by $\tilde{r}^i_i(y_i) = DD$ for

[7] Notice that any code of conduct r^i that specifies $r^i_i(y_i) = D$ for any signal y_i, and picks any map for his opponent $r^i_j(y_j) = s_j$ for all y_j, s_j would be a Nash equilibrium of G.

any y_i and any kind of mapping $\tilde{r}^i_j(y_j) = s_j$ for any s_j and y_j. Thereafter, the analysis is the same as in the one-period case.

Next, we wish to examine whether it might nevertheless be possible to have cooperation in the two period game when not only the marginal probability distribution satisfies $q < 1/5 + 4/5p$ but also strategies condition on past play. For simplicity of the exposition we analyze the case in which $p = 0$.[8] Define the code of conduct $\hat{r}^i$ for all players i as follows

$$\hat{r}^i_j(y_j) = \begin{cases} \text{tit-for-tat} & \text{if } y_j = 0, \\ DD & \text{otherwise.} \end{cases}$$

In other words, following the good signal 0 players play tit-for-tat, following the bad signal 1 players defect in both periods. If both players choose $\hat{r}^i$, they get $U_i(\hat{r}) = 10$. There are two alternative codes of conduct of interest $\tilde{r}^i, \check{r}^i$: to defect in both periods thereby ignoring the signals, or to cooperate in the first period then defect in the second period.

Consider first the code of conduct $\tilde{r}^i$ that says $\tilde{r}^i_i(y_i) = DD$ for any y_i, and any map $\tilde{r}^i_j(y_j) = s_j$ for all y_j, s_j. A player who chooses $\tilde{r}^i$ has a $1 - q$ chance of getting 6, and a q chance of getting 1 in the first period, while he gets 1 in the second period for sure. Thus, the expected utility is $U_i(\tilde{r}^i, \hat{r}^j) = 7 - 5q$. Since this is less than 10 for any q, it is never optimal the choice of $\tilde{r}^i$. Next, suppose the code of conduct $\check{r}^i$ characterized by $\check{r}^i_i(y_i) = CD$ for any y_i, and $\check{r}^i_j(y_j) = s_j$ for any y_j and s_j. Player i gets $U_i(\check{r}^i, \hat{r}^j) = 11 - 10q$. Our code of conduct $\hat{r}^i$ would be chosen over $\check{r}^i$ when $q \geq 1/10$, hence $\hat{r}$ is a Nash equilibrium of G, in fact, *strict* for all $q > 1/10$. Suppose that $q \in (1/10, 1/5)$ and $p = 0$, in the two-period game, if agents combine observation of past behavior and codes of conduct they would behave cooperatively in equilibrium. However, in the one-period interaction cooperation is not possible in equilibrium for such informative signals.

It is interesting also to see what happens in the $T < \infty$ period repeated prisoner's dilemma game with no discounting. Let us consider the time-average payoff and again concentrate on pure strategies $S_i = \{C, D\}^T$.[9] Consider the code of conduct $\hat{r}^i$ that says that both players should play the grim-strategy on the good signal, and always defect on the bad signal. We write D_t for the $t \times 1$-vector of all D entries. The code of conduct is

$$\hat{r}^i_j(y_j) = \begin{cases} \text{grim-strategy} & \text{if } y_j = 0, \\ D_T & \text{otherwise.} \end{cases}$$

This gives a payoff of exactly $U_i(\hat{r}) = 5$ (recall that $p = 0$). The optimal alternative code of conduct $\tilde{r}^i$ against $\hat{r}^i$ is to play the grim-strategy until the final period, then defect. Formally, $\tilde{r}^i_i(y_i) = (C_{T-1}, D)$ for all y_i, and for the opponent $\tilde{r}^i_j(y_j) = s_j$ for all y_j and s_j. The expected payoff would be $U_i(\tilde{r}^i, \hat{r}^j) = 1/T\,[(1-q)(5T+1) + q(T-1)]$. Hence it is optimal to adhere to $\hat{r}^i$ when $q \geq 1/(4T+2)$. The salient fact is that as $T \to \infty$ only a very tiny probability of "getting caught" is needed to sustain cooperation.

4 Repeated Games with Private Monitoring

In this section, our goal is to prove a folk-like theorem for games with private monitoring. Fudenberg and Levine (1991) consider repeated discounted games with private monitoring that are informationally connected in a way described below. They show that socially feasible payoff vectors that Pareto dominate mutual threat points are ε-sequential equilibria where ε goes to zero

[8] While studying more precise signals, they continue to be noisy.

[9] Specifically, $u_i(s) = (1/T) \sum_{t=1}^{T} g_i(a)$ where g_i is player i's stage payoff and a is the stage action profile.

as the discount factor δ tends to one. We will show that if the game is self-referential in a way that allows some chance that deviations from codes of conduct are detected (no matter how small is that chance), then this result can be strengthened from ε-sequential equilibrium to strict Nash equilibrium. We follow closely their setup.

4.1 The Stage Game

In the stage game each player $i = 1, \ldots, N$ has a (finite) action space A_i from where chooses an action a_i, and action profiles are denoted $a \in A = \times_i A_i$. Let $\Delta(A_i)$ be the probability distributions over A_i with mixed action α_i, and let $\alpha \in \Delta(A) = \times_i \Delta(A_i)$ represent mixed action profiles. Each player i has a finite private signal space Z_i with signal profiles $z \in Z = \times_i Z_i$. Given any $a \in A$, the probability of a signal profile $z \in Z$ is given by $\rho(z|a)$, and we write $\rho_i(z_i|a)$ for the marginal distribution of player i over $z_i \in Z_i$. This induces also a probability distribution for mixed actions $\alpha \in \Delta(A)$. Utility for players is denoted by $w_i : Z_i \to \mathbb{R}$ which depends only on private signal received by that player.[10] This gives rise to the expected utility function $g_i(a)$ constituting the normal form of the stage game $g_i(a) = \sum_{z_i \in Z_i} \rho(z_i|a) w_i(z_i)$. We can extend expected payoffs to $\alpha \in \Delta(A)$ in the standard way, thus $g_i(\alpha) = \sum_{a \in A} \alpha(a) g_i(a)$.

A *mutual threat point* is a payoff vector $v = (v_i, \ldots, v_N)$ for which there exists a *mutual punishment action* α such that $g_i(\alpha'_i, \alpha_{-i}) \leq v_i$ for all i, α'_i. We say a payoff vector v is *mutually punishable* if it weakly-Pareto dominates a mutual threat point. As is standard, a payoff vector v is *enforceable* if there is an α with $g(\alpha) = v$, and if for some player i and some mixed action α'_i, $g_i(\alpha'_i, \alpha_{-i}) > g_i(\alpha)$ then for some $j \neq i$ we have $\rho_j(\cdot|\alpha'_i, \alpha_{-i}) \neq \rho_j(\cdot|\alpha)$. Note that every extremal Pareto efficient payoff is enforceable.

The *enforceable mutually punishable* set V^* is the intersection of the closure of the convex hull of the payoff vectors that weakly Pareto dominate a mutual threat point and the closure of the convex hull of the enforceable payoffs. We will denote by $int(V^*)$ the interior of the set V^*. Notice that this is generally a smaller set than the socially feasible individually rational set both because there may be unenforceable actions, but also because the minmax point may not be mutually punishable.[11]

We now describe the notion of informational connectedness. Roughly this says that it is possible for players to communicate with each other even when one of them tries to prevent the communication from taking place. In a two player game there is no issue, so we give definitions in the case $N > 2$. We say that player i is *directly connected* to player $j \neq i$ despite player $k \neq j, i$ if there exists a mixed profile α and mixed action $\hat{\alpha}_i \neq \alpha_i$ for player i such that the marginal distribution of player j satisfies

$$\rho_j(\cdot|\hat{\alpha}_i, \alpha'_k, \alpha_{-(i,k)}) \neq \rho_j(\cdot|\alpha) \text{ for all } \alpha'_k.$$

In words, this condition requires that given α being played any player i's deviation will be detected by some player j regardless of player k's play. We say that i is *connected* to j if for every $k \neq i, j$ there is a sequence of players $i_1, \ldots, i_n$ with $i_1 = i, i_n = j$ and $i_p \neq k$ for any p such that player i_p is directly connected to player i_{p+1} despite player k. Intuitively, we can always find a "network" between players i and j so that the message goes through no matter what other single player tries

[10] We may include the players own action in his signal if we wish.

[11] Fudenberg and Levine (1991) show only that V^* contains approximate equilibria leaving open the question of when the larger socially feasible individually rational set might have this property. They construct approximate equilibria using mutual punishment, hence there is no effort to punish the player who deviates. This is necessary because they do not impose informational restrictions, of the type imposed in Fudenberg et al. (1994), sufficient to guarantee that it is possible to determine who deviated. With those restrictions it is likely that their methods would yield a stronger result. As this is a limitation of the original result, we do not pursue the issue here.

to do. The game is *informationally connected* if there are only two players, or if every player is connected to every other player.

4.2 The Repeated Game

We now consider the T repeated game with discounting, where we allow both T finite and $T = \infty$. A history for player i at time t is a sequence $h_i^t = (a_i^1, z_i^1, \ldots, a_i^t, z_i^t)$ while $h_i^0 = \emptyset$ is the null history. The set of all t-length private histories for player i is denoted by $H_i^t = (A_i \times Z_i)^t$ and the set of all private histories for player i by $H_i = \bigcup_t H_i^t$. A behavior strategy for player i is a map $\sigma_i : H_i \to \Delta(A_i)$. We write σ for the profile of behavior strategies. Players have common discount factor $\delta \in (0, 1)$. For some δ we let $u_i(\sigma; \delta, T)$ denote expected average present value for the game repeated T periods.

Combining Theorems 3.1 and 4.1 from Fudenberg and Levine (1991) we have the following theorem:

Theorem 4.1 (Fudenberg and Levine (1991))**.** *In an informationally connected game if $v \in V^*$ then there exists a sequence of discount factors $\delta_n \to 1$, non-negative numbers $\varepsilon_n \to 0$ and strategy profiles σ_n such that σ_n is an ε_n-Nash equilibrium*[12] *for δ_n and $u(\sigma_n; \delta_n, \infty) \to v$.*

We will also use the fact that their Lemma A.2 together with their construction in the proof of Theorem 4.1 implies that it is possible to build a *communication phase* with length L such that the following holds.

Lemma 4.1 (Fudenberg and Levine (1991))**.** *For any $\beta \in (0, 1)$ there exists a pair of strategies σ_i, σ_i' and for each player $j \neq i$ a test $\mathcal{Z}_j \subset \{(z_j^1, \ldots, z_j^L)\}$ such that for any player $k \neq i, j$ and strategy σ_k' by player k, under $(\sigma_i, \sigma_k', \sigma_{-(i,k)})$ we have $Pr((z_{-(i,k)}^1, \ldots, z_{-(i,k)}^L) \in \mathcal{Z}_{-(i,k)}) \leq 1 - \beta$, and under $(\sigma_i', \sigma_k', \sigma_{-(i,k)})$ we have $Pr((z_{-(i,k)}^1, \ldots, z_{-(i,k)}^L) \in \mathcal{Z}_{-(i,k)}) \geq \beta$.*

This says that a player can "communicate" to the entire group by using his actions whether or not someone has deviated. In fact, such communication between players is guaranteed by the assumption of information connectedness.

4.3 The Finitely Repeated Self-Referential Game

In the self-referential case it is convenient to work with finite versions of the repeated game. The T-discrete version of the game has finite time horizon T and players have access each period to independent randomization devices that provide a uniform over T different outcomes. The self-referential T-discrete game Γ^T consists of signal spaces Y_i for each player i, codes of conduct space R_T, and the signal probabilities are given by $\pi_T(\cdot|r)$. The self-referential game G is said to E, D *permit detection* where constants E, D satisfy $E, D \in [0, 1]$ if for every player i there exists a player j and a nonempty set $\bar{Y}_j \subset Y_j$ such that for any profile code of conduct $r \in R$, any signal $\bar{y}_j \in \bar{Y}_j$, and any $\tilde{r}^i \neq r^i$ we have $\pi_j(\bar{y}_j|\tilde{r}^i, r^{-i}) - \pi_j(\bar{y}_j|r) \geq D$ and $\pi_j(\bar{y}_j|r) \leq E$.

Next we state the main result of the paper:

Theorem 4.2. *If $int(V^*) \neq \emptyset$, the game is informationally connected, for some $E \geq 0, D > 0$ the self-referential T discrete versions E, D permits detection, and $v \in int(V^*)$ then for any ε_0 there exists a sufficiently large discount factor δ, a discretization T and strict Nash equilibrium codes of conduct $\hat{r} \in R$ such that for all i, $|v_i - U_i(\hat{r})| \leq \varepsilon_0$.*

[12] Fudenberg and Levine (1991) prove the stronger result that σ_n is an ε_n-sequential equilibrium which means also that losses from time t deviations measured in time t average present value, and not merely time t average present value, are no bigger than ε_n. As we do not need it, we omit the extra definitions required to state the stronger result.

Proof. By choosing T^* large enough for given δ it follows that ε_0-Nash equilibria of the infinitely repeated game are $2\varepsilon_0$-Nash equilibria of the discretized game Γ^{T^*}, so Theorem 4.1 applies to Γ^{T^*}. Thus, Theorem 4.1 implies that for all sufficiently large δ^* we can find a sequence $\{\delta_n\}$ with $\delta_n \geq \delta^*$ and corresponding sequence of times $\{T_n\}$ together with strategies $\bar{s}^0, \bar{s}^1, \ldots, \bar{s}^N$such that these are all $\frac{\varepsilon_0}{2}$-Nash equilibria, that for each player i, $|u_i(\bar{s}^0) - v_i| < \varepsilon_0/2$, $u_i(\bar{s}^0) - u_i(\bar{s}^i) \geq 2\sqrt[3]{\varepsilon_0}$ for each such strategies $\bar{s}^i$, and $|u_j(\bar{s}^i) - u_j(\bar{s}^0)| < \varepsilon_0/2$ for each player j and strategy $\bar{s}^i$.

For each pair of players i, j we construct a pair of strategies s_i^0, s_i^j as follows. We begin the game with a series of communication phases $\{C_j\}_{j=1}^N$ where actions are used to communicate the deviator's identity. We go through the players $j = 1, \ldots, N$ in order each phase C_j lasting L periods. In the first jth phase C_j, the player $i \neq j$ who is able to detect deviations by player j has two strategies $\hat{s}_i^j, \hat{s}_i^{j'}$ and players $k \neq i, j$ have a strategy $\hat{s}_k^j$ from Lemma 4.1. During the phase C_j let s_i^0 be the strategy such that player i plays the L truncation of $\hat{s}_i^j$, alternatively we define the strategy s_i^j such that he picks the L truncation of $\hat{s}_i^{j'}$ thereby sending the message that player j has deviated. While the remaining players k play the L truncation of $\hat{s}_k^j$.

At the end of these $N \times L$ periods of communication we specify that in the strategy profile s^0 each player j conducts the test $\mathcal{Z}_j \subset \{(z_j^1, \ldots, z_j^L)\}$ for each C_j in Lemma 4.1 to statistically check who has sent a signal for a given level of precision β. If it indicates that exactly one player i has sent a signal or that exactly two players i, j sent a signal where i reports that j has deviated then he plays his part of the equilibrium punishment strategy $\bar{s}^j$ punishing player j. In all the other cases, he picks his part from the equilibrium strategy $\bar{s}^0$. Under any of the strategies s_i^0, s_i^j for player i in each C_j, from Lemma 4.1 by choosing sufficiently large L the probability β that all players agree that a single player i sent a signal (since in fact at most one player has actually sent a signal) or that no signal was sent may be as close to 1 as we wish. In particular we may choose β close enough to 1 that play following disagreement or agreement on more than one player sending a signal has no more than an $\varepsilon_0/4$ effect on payoffs.

Observe that the choice of the length of communication phases L does not depend on δ_n, T_n, so we may choose δ_n and T_n large enough that nothing that happens in the communications phase $\{C_j\}_{j=1}^N$ makes more than a $\varepsilon_0/4$ difference to payoffs.

From the previous construction we have found a pair of strategies s_i^0, s_i^j for each pair of players i, j such that the strategy profile s^0 is an ε_0-Nash equilibrium for δ^*, T^*. For any player j, the corresponding strategy profile s^j is an ε_0-Nash equilibrium as well. Note that for any punishment strategy s^j it follows that $|u_i(s^j) - u_i(s^0)| < \varepsilon_0$ for all players i. Furthermore, we have constructed each s^i so that $u_i(s^0) - u_i(s^i) \geq \sqrt[3]{\varepsilon_0}$ for any punished player i and in the equilibrium strategy s^0 players approximately obtain v_i. Given this construction, by the bounds in Theorem 5.1 there exists a strict Nash equilibrium codes of conduct profile $\hat{r}$ such that players get nearly what they get in the approximate equilibrium s^0, that is, $|v_i - U_i(\hat{r})| \leq \varepsilon_0$. □

It is worth understanding the idea behind this construction. The signals concerning the code of conduct are weak (because we have fixed E). Hence we cannot use very strong nor yet mutual punishment to prevent deviations from the code of conduct. On the other hand the weak signals do provide information about who violated the code of conduct. Hence we construct a family of approximate equilibria in the original game with similar payoffs each representing a small punishment for a particular player. If there is evidence that a particular player violated the proposed code of conduct then everyone switch to the approximate punishment equilibrium against that player. Players then face the choice: follow the code of conduct and forego a small gain to deviating (since the code of conduct calls for an approximate equilibrium of the original game), or violate it and get a small punishment with small probability. We then calibrate the parameters so that the expected cost of the punishment is greater than the small gain to deviating. The approximate equilibria themselves following Theorem 4.1 have a very different structure: in the approximate equilibria very precise information is accumulated on how players have played and a mutual punishment is used, but so infrequently on the equilibrium path it has little cost.

5 Conclusion

The standard world of economic theory is one of perfect liars – a world where scammers have no difficulty passing themselves off as businessmen. In practice social norms are complicated and there is some chance that a player will inadvertently reveal his intention to violate a social norm through mannerisms or other indications of lying. Here we investigate a simple model in which this is the case. Our setting is that of self-referential games, which allows the possibility of observing directly opponents' intentions. We characterized the self-referential nature of this class of games by defining codes of conduct which represent agreement between players that even have different roles.

In practice the probability of detection is not likely to be perfect, so we focus on the case where the detection probability is small. The key idea is that a little chance of detection can go a long way. Small probabilities of detecting deviation from a code of conduct allow us to sustain approximate equilibria as strict equilibria of the self-referential version of the game. An illustrative, but far more important, application of this result is a discounted strict Nash folk-like theorem for repeated games with private monitoring. We conclude that approximate equilibria can be sustained as "real" equilibria when there is a chance of detecting violations of codes of conduct.

We have assumed that players can observe signals about the opponents' intentions only at the beginning of the game as is standard in the literature. Yet if the base game is a repeated game it is natural to wonder whether our folk theorem holds when signals arrive over time. Block (2013) explores the case where agents sequentially learn the opponents' intentions by privately observing their signal at some period of the game.[13] One key insight is that a folk theorem results for repeated games with perfect monitoring regardless of the period at which signals can be observed, and the earlier agents see the signal the lower is the probability of detection required to construct the equilibrium in the self-referential game. We conjecture that our folk theorem would hold if agents receive signals in the course of the game but one might expect to have the probability of detection depending on the period at which signals arrive.

Appendix

We assume that in the base game all players have access to N individual randomizing devices $\Theta = \{\theta^1, \dots, \theta^N\}$ each of which has an independent probability ε_R of an outcome we call *punishment*, θ_p. Suppose that s^0 is an ε_0-Nash equilibrium giving utility $u_i(s^0)$ for each i. For any $s \in S$ and strategies s^i_j for any pair of players $i \neq j$ suppose that $s^i = (s^i_j, s_{-j})$ are ε_1-Nash equilibria. Define $P_i = u_i(s^0) - u_i(s^i)$. We assume that $P_i \geq \underline{P} \geq 0$ and for some $\varepsilon_p \geq 0$ that $|u_j(s^i) - u_j(s^0)| \leq \varepsilon_p$. Let $\underline{u} = \min_{i,s} u_i(s)$ and $\bar{u} = \max_{i,s} u_i(s)$ be the lowest and highest payoffs, respectively. Define $\varepsilon = \varepsilon_0 + (N + \bar{u} - \underline{u})(\varepsilon_1 + \varepsilon_p)E$, and $K = \max\{(N + \bar{u} - \underline{u})3N^4(1 + \bar{u} - \underline{u}), (N^4(\bar{u} - \underline{u}) + 1)(\bar{u} - \underline{u})\}$.

Theorem 5.1. *If $(D(\underline{P} - \varepsilon_1))^2 > 4K\varepsilon$, then there is an ε_R and strict Nash equilibrium codes of conduct $\hat{r} \in R$ such that for all players i, $|u_i(s^0) - U_i(\hat{r})| \leq \varepsilon + D(\underline{P} - \varepsilon_1) - \sqrt{(D(\underline{P} - \varepsilon_1))^2 - 4K\varepsilon}$.*

[13] To highlight the role of the timing a single signal is assumed to be observed at a predetermined period, however, the interpretation is that players aggregate information during the game and then they make use of this information. It also accommodates the relevant case where players receive signals over time in a repeated game.

Proof. There are $|\Theta|^{|I|} = N^2$ independent randomization devices in operations. From $P_\theta(\theta_p = 0) = (1-\varepsilon_R)^{N^2}$ and $P_\theta(\theta_p = 1) = N^2\varepsilon_R(1-\varepsilon_R)^{N^2-1}$ we find $P_\theta(\theta_p \geq 2) \leq N^4\varepsilon_R^2$.

Take $\hat{r}^i$ such that: for all players i, if $\bar{y}_i \in \bar{Y}_i$ and $\theta_p \geq 1$ play $\hat{r}^i_i(y_i) = s^j_i$ and $\hat{r}^i_j(y_j) = s^0_j$ for any $y_j \in Y_j$ and all players $j \neq i$, otherwise play $\hat{r}^i_i(y_i) = s^0_i$ for all $y_i \notin \bar{Y}_i$ and for all $j \neq i$ choose $\hat{r}^i_j(y_j) = s^0_j$ for any $y_j \notin \bar{Y}_j$. The following mutually exclusive events can occur to player i when all players $j \neq i$ choose $\hat{r}^j$, but he chooses $\tilde{r}^i$ defined below and $\tilde{r}^i_j = \hat{r}^i_j$ for all $j \neq i$: (1) Nobody is punished: if player i chooses $\hat{r}^i$ he gets $u_i(s^0)$, while if i chooses $\tilde{r}^i_i \neq s^0_i$ he gets at most $u_i(s^0)+\varepsilon_0$, (2) Player j is the only player punished: by following $\hat{r}^i$ i gets $u_i(s^j)$, if i chooses $\tilde{r}^i_i(\bar{y}_i) \neq s^j_i$, he gets at most $u_i(s^j)+\varepsilon_1$, and (3) Two or more players are punished: if $\hat{r}^i$ is followed player i he gets at worst $\underline{u}$, if i deviates while choosing $\tilde{r}^i(y_i) = \tilde{s}_i$ with $\tilde{s}_i \in S_i$ and $\tilde{s}_i \neq s^j_i \neq s^0_j$ he gets at most $\bar{u}$.

We can bound expected payoffs $U_i(\hat{r}) \leq \overline{U}_i(\hat{r}) = u_i(s^0) + (1-(1-E)^N)\left[\varepsilon_p + N^4\varepsilon_R^2(\bar{u}-\underline{u})\right]$ and $U_i(\hat{r}) \geq \underline{U}_i(\hat{r}) = u_i(s^0) - (1-(1-E)^N)\left[\varepsilon_p + N^4\varepsilon_R^2(\bar{u}-\underline{u})\right] - \pi_j(\bar{y}_j|\hat{r})\varepsilon_R P_i$. Suppose player i chooses the alternative code of conduct $\tilde{r}^i$, he gets $U_i(\tilde{r}^i, \hat{r}^{-i}) \leq \overline{U}_i(\tilde{r}^i, \hat{r}^{-i})$ where

$$\overline{U}_i(\tilde{r}^i, \hat{r}^{-i}) = u_i(s^0)+\varepsilon_0+(1-(1-E)^N)\left[\varepsilon_1 + N^4\varepsilon_R^2(\bar{u}-\underline{u})\right] - \left[(\pi_j(\bar{y}_j|\hat{r})+D)\varepsilon_R - N^4\varepsilon_R^2\right](P_i+\varepsilon_1).$$

Consequently, the gain to choosing $\tilde{r}^i$ is $U_i(\tilde{r}^i, \hat{r}^{-i}) - U_i(\hat{r})$ and bounded by

$$\begin{aligned}&\varepsilon_0 + (1-(1-E)^N)\left[\varepsilon_1+\varepsilon_p+2N^4\varepsilon_R^2(\bar{u}-\underline{u})\right] + \pi_j(\bar{y}_j|\hat{r})\varepsilon_R P_i - \left[(\pi_j(\bar{y}_j|\hat{r})+D)\varepsilon_R - N^4\varepsilon_R^2\right](P_i-\varepsilon_1)\\ &\quad\leq \varepsilon_0 + (N+\bar{u}-\underline{u})E\left[\varepsilon_1+\varepsilon_p+3N^4\varepsilon_R^2(1+\bar{u}-\underline{u})\right] - D\varepsilon_R(\underline{P}-\varepsilon_1) \leq \varepsilon + K\varepsilon_R^2 - D\varepsilon_R(\underline{P}-\varepsilon_1).\end{aligned}$$

Hence if $D\varepsilon_R(\underline{P}-\varepsilon_1) > \varepsilon + K\varepsilon_R^2$ then there is a strict Nash equilibrium with

$$\left|u_i(s^0) - U_i(\hat{r})\right| \leq NE\varepsilon_p + (N^4(\bar{u}-\underline{u})+1)(\bar{u}-\underline{u})\varepsilon_R \leq \varepsilon + 2K\varepsilon_R.$$

We conclude by solving the inequality for $\varepsilon_R = [D(\underline{P}-\varepsilon_1) \pm \sqrt{(D(\underline{P}-\varepsilon_1))^2 - 4K\varepsilon}]/2K$, which gives two real roots since $(D(\underline{P}-\varepsilon_1))^2 > 4K\varepsilon$, implying the existence of an ε_R for which $\hat{r}$ is a strict Nash equilibrium of the self-referential game. Plugging the lower root into the inequality for the utility difference $\left|u_i(s^0) - U_i(\hat{r})\right|$ gives the remainder of the result. □

References

Bachi, B., S. Ghosh, and Z. Neeman (2014). Communication and deception in 2-player games. *Working paper*.

Block, J. I. (2013). Timing and codes of conduct. *Working paper*.

Charness, G. and M. Dufwenberg (2006). Promises and partnership. *Econometrica 74*(6), 1579–1601.

Compte, O. (1998). Communication in repeated games with imperfect private monitoring. *Econometrica 66*(3), 597–626.

Ely, J. C., J. Hörner, and W. Olszewski (2005). Belief-free equilibria in repeated games. *Econometrica 73*(2), 377–415.

Frank, R. H. (1988). *Passions Within Reason: The Strategic Role of the Emotions.* Norton : New York.

Fudenberg, D. and D. K. Levine (1991). An approximate folk theorem with imperfect private information. *Journal of Economic Theory 54*(1), 26–47.

Fudenberg, D. and D. K. Levine (2007). The nash-threats folk theorem with communication and approximate common knowledge in two player games. *Journal of Economic Theory 132*(1), 461 – 473.

Fudenberg, D., D. K. Levine, and E. Maskin (1994). The folk theorem with imperfect public information. *Econometrica 62*(5), 997–1039.

Gauthier, D. P. (1986). *Morals by agreement.* Oxford University Press : New York.

Hirshleifer, J. (1987). *On the emotions as guarantors of threats and promises.* Cambridge: MIT Press.

Hörner, J. and W. Olszewski (2006). The folk theorem for games with private almost-perfect monitoring. *Econometrica 74*(6), 1499–1544.

Kalai, A. T., E. Kalai, E. Lehrer, and D. Samet (2010). A commitment folk theorem. *Games and Economic Behavior 69*(1), 127–137.

Kandori, M. and H. Matsushima (1998). Private observation, communication and collusion. *Econometrica 66*(3), 627–652.

Lehrer, E. (1990). Nash equilibria of n-player repeated games with semi-standard information. *International Journal of Game Theory 19*(2), 191–217.

Levine, D. K. and W. Pesendorfer (2007). The evolution of cooperation through imitation. *Games and Economic Behavior 58*(2), 293–315.

Matsushima, H. (2004). Repeated games with private monitoring: two players. *Econometrica 72*(3), 823–852.

Peters, M. (2015). Reciprocal contracting. *Journal of Economic Theory 158, Part A*(0), 102–126.

Peters, M. and B. Szentes (2012). Definable and contractible contracts. *Econometrica 80*(1), 363–411.

Renou, L. and T. Tomala (2013). Approximate implementation in markovian environments. *HEC Paris Research Paper No. ECO/SCD-2013-1015*.

Sally, D. (1995). Conversation and cooperation in social dilemmas a meta-analysis of experiments from 1958 to 1992. *Rationality and society 7*(1), 58–92.

Schelling, T. C. (1978). Altruism, meanness, and other potentially strategic behaviors. *American Economic Review 68*(2), 229–30.

Sugaya, T. (2011). Folk theorem in repeated games with private monitoring. *Economic Theory Center Working Paper* (011-2011).

Szentes, B. (2015). Contractible contracts in common agency problems. *Review of Economic Studies 82*(1), 391–422.

Tennenholtz, M. (2004). Program equilibrium. *Games and Economic Behavior 49*(2), 363–373.

Zuckerman, M., B. M. DePaulo, and R. Rosenthal (1981). Verbal and nonverbal communication of deception. *Advances in experimental social psychology 14*, 1–59.