Penalty-card strategies in repeated games

Wojciech Olszewski and Mikhail Safronov*

February 12, 2015

Abstract

We study a class of what we call penalty-card strategies in repeated games of incomplete information. The idea is that a player who plays an action resulting in a low expected payoff of other players may obtain a penalty card. If a player obtains a limit number of cards, she is going on suspension; during the period of suspension she must play an action resulting in a high expected payoff of other players. Any deviation while being on suspension results in a breakdown of cooperation.

We show that if players' privately known types are i.i.d., or more generally evolve according to a Markov chain, then under some mild conditions on the stage game, the outcomes that maximize the aggregate payoff of all players can be attained in penalty-card strategies for the discount factor tending to 1.

Penalty-card strategies have several useful features, e.g., players condition their actions only on a simple statistics containing all necessary information regarding the past play, and can be viewed as a positive model of playing repeated games with incomplete information.

1 Introduction

The models of repeated games of incomplete information have a wide range of applications. They include: (a) oligopoly markets in which firms privately know their costs, (b) repeated auctions in which bidders privately know their valuations; (c) partnership games in which the effort that can be exerted depends on other duties that partners must perform, or (d) favor exchange when a person in need does not know if others can help.

In the existing literature, repeated games of incomplete information have been analyzed by means of two kinds of strategies: (a) simple and intuitive strategies that allow to obtain only limited results, or (b) strategies that allow to attain a wider range of payoffs, but are less intuitive and more involved, or have been "tailored" with the objective of attaining particular payoffs.

The aim of this paper is to introduce a class of as we believe simple and intuitive strategies, which in addition support a wide range of outcomes. We study a class of what we call *penalty-card strategy profiles.*

*Department of Economics, Northwestern University, 2001 Sheridan Rd., Evanston IL 60208-2600; E-mails: wo@northwestern.edu and MikhailSafronov2014@u.northwestern.edu, respectively

The idea is that a player who plays an action which "hurts" other players obtains a (tacit) penalty card. One card has no immediate effect on the future play. But if a player obtains a limit number of cards, she goes on suspension, which means that for some number of periods she has to take actions that other players "want" her to play. Detectable deviations result in a breakdown of cooperation.

Penalty-card strategies resemble what we observe in numerous settings in practice; in fact, the idea comes from penalty cards as used in sport games, and the concept of warning as used in numerous everyday interactions. Players condition their actions only on simple statistics containing all necessary information regarding the past play. The strategy profiles also have a number of other useful features.

We show that if players' privately known types are i.i.d., or more generally evolve according to a Markov chain, then under some mild conditions on the stage game, the efficient outcome (that is, the outcome that maximizes the aggregate payoff of all players) can be attained in penalty-card strategies for the discount factor tending to 1. The fact that the efficient outcome can be attained is roughly intuitive. When the discount factor becomes closer to 1, and players assign higher weights to future payoffs, one can allow a larger limit number of penalty cards. So, players go on inefficient suspension less frequently, and the action profile that maximizes the aggregate payoff is played more frequently.

The fact that the penalty-card strategies supporting the efficient outcome are incentive compatible requires a careful design of the transition probability in the structure of penalty cards in response to players' actions. We design this transition probability by imitating the d'Aspremont and Gerard-Varet (1979) mechanism. That is, the structure of penalty cards changes in the way that players internalize the current payoffs of their opponents by the effect that their actions have on their own continuation payoff.

Despite the fact that current actions in penalty-card strategies are contingent only on a simple statistics regarding the past play, the number of incentive constraints which must be checked is still large, and their form is quite complicated. We make the analysis more tractable by studying the incentive constraints only for the discount factor δ tending to 1. This allows for omitting all expressions of order $o(1-\delta)$, which makes the form of constraints simpler, and enables us to derive explicit formulas for the repeated-game payoffs.

An additional difficulty arises when players' types are persistent or correlated over time, because players' current actions may reveal to their opponents information about their future types, which may affect their opponents' future actions. These "signalling" and "ratcheting" effects might suggest that players must condition their strategies on the previous actions of their opponents not only through penalty cards. Strategies contingent on previous actions are, however, not necessary. We prescribe the changes in the structure of penalty cards in the way that players not only internalize the effect of their current actions on the current payoffs of the opponents, but also internalize the expected effect of their current actions on the future payoffs.

In the present paper, we show our results for stage games with finite number of actions and types. However, as we show in a companion paper Olszewski and Safronov (2015), similar results hold for a number of games with infinite action space, for example, in repeated auctions and a repeated version of Spulber's (1995) oligopoly game, studied in numerous existing papers. We also show in the companion paper that in many applications, including versions of (a)-(d), efficiency can be attained in particularly simple and intuitive strategies.

Related literature

Fudenberg, Levine and Maskin (1994) show a folk theorem for a family of repeated games with finite numbers of actions and types, in which players have i.i.d. types. The focus of their paper is entirely on the payoffs that can be attained in equilibrium, not on the strategies attaining these payoffs.

More recent research on the topic was initiated by papers on repeated duopoly and repeated auctions. In Athey and Bagwell (2001), followed up by Athey, Bagwell and Sanchirico (2004), firms play a repeated version of Spulber's (1995) oligopoly game with an infinite number of actions, and each firm is privately informed of its cost of production. This cost follows an i.i.d. process. Among other results, they show that the efficient payoff vector can be attained in the two-firm case when the discount factor exceeds some cutoff level. Hörner and Jamison (2007) generalize this result to an arbitrary number of firms. Athey and Bagwell (2008) extend their 2001 model to the more realistic case in which the firms' costs are more persistent (more precisely, they follow a Markov process). They construct an equilibrium which depends on the firms' costs, and which attains a more efficient payoff vector than the best equilibrium in strategies which are independent of the firms' costs. This more efficient payoff is, however, not efficient. They also construct an efficient equilibrium, but only in the two-firm case. Finally, Escobar and Toikka (2010) show that the efficient payoff vector can be attained in suitably modified review strategies; they even prove that any Pareto-efficient payoff vector above a stationary minmax vector can be attained for a generic class of games.

Athey and Bagwell (2001), Athey, Bagwell and Sanchirico (2004), and Athey and Bagwell (2008) construct intuitive and simple equilibrium strategies, but obtain only limited efficiency results.[1],[2] Hörner and Jamison's (2007) result is general and obtained under weak assumptions regarding the observability of actions, but the strategies are quite involved and carefully "tailored" for obtaining particular payoffs. In addition, all these authors study only the Spulber's oligopoly game, and except Athey and Bagwell (2008) these papers assume i.i.d. types. The review strategies[3] used by Escobar and Toikka are intuitive, and deliver general results. However, their complete strategies are not entirely explicit. Contingent on some histories, they are defined by a fixed-point argument, although the chance of reaching such a history is rather low.

Skrzypacz and Hopenhayn (2004) Blume and Heidhues (2002) study collusion in repeated auctions, and show that players can obtain a higher equilibrium total payoff that in the repetition of stage-game Nash

[1] The equilibria constructed in Athey and Bagwell (2008) resemble our "one-penalty card" strategy profiles. So, not surprisingly, they achieve only partial, but not full efficiency.

[2] In addition, Athey and Bagwell construction requires the assumption that firms can split the market unequally, in the way they wish, when they charge equal prices. As we show in the working paper, the penalty-card strategies attain efficiency not only for any number of firms, but also when the market splits equally among the firms which charge the lowest prize.

[3] Review strategies were initially studied by Radner (1985) in a repeated moral hazard game. In the case of i.i.d. costs, the idea behind Escobar and Toikka's equilibria are also closely related to the linking mechanism from Jackson and Sonnenschein (2007).

equilibrium, and in the bid rotation scheme which in every period appoints a winning bidder, making in exchange other bidders more likely to be winners in future periods. Allowing for mediated communication, Aoyagi (2007) shows that efficiency can be attained for a large class of repeated-auction settings. He obtains these results by modifying self-generation techniques. The strategies that attain efficient outcomes, as well as the strategies studied in Aoyagi (2003), share numerous features with our penalty-card strategies. Namely, the stage-auction winners are excluded for a number of periods, and the winners in the following periods are efficiently appointed from the set recent losers.

Another class of related strategies studied in the existing literature, but in quite different settings is the class of chip strategies. According to these strategies, each player is initially endowed with a certain number of chips; a player who plays an action such that her opponents' payoff is high obtains from them a chip, or gives them a chip if her opponents' payoff is low. Möbius (2001) and Hauser and Hopenhayn (2008) analyze a model of voluntary favor exchange between two players. Favor opportunities arrive according to a Poisson process, and the benefit of receiving a favor exceeds the cost of providing it. Mobius identifies conditions under which chip strategies constitute an equilibrium. For any given discount factor, the equilibria in chip strategies cannot obviously be fully efficient, because incentive compatibility imposes a limit on the number of chips that can be used. Hauser and Hopenhayn suggest two improvements to chip strategies that enhance the efficiency of equilibria: exchanging chips at different rates (i.e., one favor today is not necessarily equivalent to one favor in the future), and appreciation and depreciation of chips. Solving the model numerically, they demonstrate that for a large set of parameter values the efficiency gains are quite large. Abdulkadirouglu and Bagwell (2012) analyze the chip mechanism in a discrete version of Möbius' model of favors. They identify the optimal limit number of chips given a discount factor, and compare this optimal chip mechanism with a more sophisticated favor-exchange relationship in which the size of a favor owed may decline over time. For any given discount factor, the equilibria in chip strategies cannot obviously be fully efficient, because incentive compatibility imposes a limit on the number of chips that can be used. They provide no efficiency result (explicitly or implicitly) when the discount factor converges to 1.[4]

We do not pretend to improve on the existing literature in terms of generality of efficiency results. Indeed, to generalize Fudenberg, Levin and Maskin's, and Escobar and Toikka's result to a generic class of games, one would have to assume that types evolve according to a non-Markov process, or relax the assumption of perfect monitoring of actions. We view our main contribution as describing intuitive strategies, which provide a positive model of playing repeated games with incomplete information. These strategies, under relatively mild conditions, attain efficiency in all settings studied in the existing literature, that is, for i.i.d. and Markov types, and for games with finite action and type spaces, and even in some applications with infinite action spaces.[5] One may also find interesting the relation to the d'Aspremont and Gerard-Varet

[4] One should also mention Athey and Miller (2007) who look at similar debt strategies in a model of repeated trade with hidden valuations.

[5] We also conjecture, but we have not proved formally, that one can obtain efficiency results in repeated auctions without

(1979) mechanism. This mechanism has several useful properties, and they are inherited by our repeated game strategies. We will emphasize some other advantages of penalty-card strategies in Section 7.

In Section 2, we introduce the model, and present verbally penalty-card strategies. In Section 3, we state the result, and describe the main idea behind our construction of equilibria, focusing on the i.i.d. case. Section 4 contains the detailed exposition of "efficient" penalty-card strategies in this case. Sections 5 and 6 are devoted to the proof that these strategies satisfy equilibrium conditions, and approach the efficient outcome for the discount factor tending to 1. The proof in the i.i.d. case is relatively simple, exhibits all basic ideas, avoiding more delicate issues specific to the Markovian case. We state the result, and point out the key modification in the construction of equilibria required in this case in Section 8, but postpone the detailed proof to Appendix. Finally, we elaborate in Section 7 on the advantages of penalty-card strategies.

2 Preliminaries

2.1 Model

Consider a normal-form game G with I players, numbered by $i = 1, ..., I$. Let A_i and Θ_i be finite sets of actions and (privately observed) types, respectively, of player i. Let $u_i(\theta_i, a)$ be the payoff of player i. We make some, mild assumptions on the payoffs. These assumptions will be better understood when penalty-card strategies are defined, so we will introduce them later.

We study a repeated game in which players play stage game G in periods $t = 1, 2, ...$, and discount future payoffs at a common rate δ; it is convenient to denote $1 - \delta$ by ε. Actions are publicly observed at the end of each period. In the repeated game, players are allowed to communicate by sending at the beginning of each period simultaneous, publicly observed cheap-talk messages regarding their types. We assume that the message space of each player i coincides with the type space Θ_i. Players also have access to a public randomization device, i.e., they observe the realization of a random variable distributed uniformly on interval $[0, 1]$. The timing of events in each period is as follows: (a) players privately observe their types; (b) they send public cheap-talk messages regarding their types; (c) players take publicly observed actions; (d) they observe a realization of public random device.

In the main text, we assume that players' types are i.i.d. according to distributions η_i, $i = 1, ..., I$. In Appendix, we generalize the results from the main text to players' types which are still independently distributed, but evolve over time according to homogeneous, aperiodic irreducible Markov chains. That is, if the current-period type profile is $\theta = (\theta_1, ..., \theta_I) \in \Theta = \prod_{i=1}^{I} \Theta_i$, the next-period type profile will be θ' with (transition) probability $\eta_{\theta,\theta'}$, and for every pair of type profiles $\theta, \theta' \in \Theta$, there exists a t such that if the current-period type profile is θ, the type profile in period t will be θ' with positive probability. By ergodic theorem, every such process has a limiting type distribution η, and independent of the initial type profile,

communication, improving on Aoyagi (2007).

the distribution of types at time t converges as $t \to \infty$ to the limiting distribution at an exponential rate.

All other elements of the model, that is, histories, repeated-game strategies and payoffs are defined in the standard manner.

2.2 Description of penalty-card strategies

We study the following class of strategies:

- A *penalty-card strategy profile* has two phases: a *cooperation phase*; and a *joint-penalty phase*.
- In the cooperation phase, some players are on *suspension*, and other players are *active*. Initially, all players are active.
- Actions prescribed for both active players and players on suspension depend only on the set of active players.

(Typically, the prescribed actions "reward" active players and "penalize" players on suspension.)

- Each active player holds a certain number of *penalty cards*. A player may obtain another penalty card, or some of the exisitng cards may be annulled.
- Any player can collect only up to a certain number of penalty cards. If a player reaches this limit number, the player goes on suspension for a certain, possibly random number of periods.
- Players who come back from suspension become active. This happens when the prescribed suspension comes to its end, independent of the actions of players.
- The chance of obtaining another penalty card by some player, or of annulling some existing cards is determined by the current penalty-card structure and the actions of active players;
- The play stays in the cooperation phase until some player takes an action which is not prescribed for any private information.
- Any unprescribed action triggers the joint-penalty phase. Once the joint penalty phase is triggered the play remains in this phase forever.

The definition is motivated by the concept of penalty cards as used in sport games, or the concept of warning as used in numerous everyday interactions, and is aimed to adjust these concepts in the possibly simplest manner to the repeated-game setting.

We prove our results using more specific penalty-card strategies. For example, only one active player will hold penalty cards. One may also consider slightly more general classes strategies with similar features.

For example, a unilateral deviation could trigger a player-specific penalty phase, instead of any deviation triggering a joint-penalty phase.

The penalty-card strategies can also be viewed as a specific debt contract in which players holding penalty cards are borrowers, and their debt is jointly owned by other players. Players on suspension are (possibly temporarily) excluded from the credit market, and forced to repay their debt, and joint-penalty phase may be interpreted as a credit-market failure.

2.3 Assumptions on stage game

We can now present the assumptions that we impose throughout the main text on the stage game.[6] For any set of players $R \subset \{1, ..., I\}$ and their type profile $\theta_R \in \Theta_R = \prod_{i \in R} \Theta_i$, denote by $a(\theta_R)$ a *R-efficient action profile*, that is, an action profile that maximizes the total payoff of the players from R.[7] Let $v_R^i = E_\theta(u_i(\theta_i, a(\theta_R)))$ for $i \notin R$ denote the expected stage-game payoff of a player i who is not in R, when players take the R-efficient action profile. Similarly, let w_R^i for $i \in R$ denote the expected stage-game payoff of player i who is a member of R, when players take the R-efficient action profile.

Assumption I: For any $i = 1, ..., n$ and R such that $i \notin R$,

$$v_R^i < w_{R \cup \{i\}}^i.$$

Assumption II: For any $i = 1, ..., n$ and R such that $i \in R$,

$$\frac{1}{|R| - 1} \sum_{i \neq j \in R} w_{R-\{j\}}^i > v_{R-\{i\}}^i.$$

To interpret these assumptions suppose that in the cooperation phase of a penalty-card strategy profile, players play R-efficient action profile, for R being the set of currently active players. Then, our assumptions say that every player prefers being active to being on suspension (for any subset of other active players), and prefers on average when another player is on suspension to being on suspension herself.

Finally, we assume that

Assumption III: The incomplete information stage game has an equilibrium in which the payoff of every player i is lower than w_R^i for $R = \{1, ..., I\}$.

We will call the equilibrium described in assumption III *bad equilibrium*. The existence of stage-game equilibria for general games can be established by a simple fixed-point argument. However, stage-game

[6] That is, these assumptions apply to the i.i.d. case. In the Markovian case studied in Appendix, the assumptions will be very similar but slightly stronger.

[7] Pick an arbitrary maximizer when there exist more than one.

equilibria may not satisfy assumption III. There may exist no equilibrium satisfying assumption III even in complete information games with degenerate type spaces.

However, bad equilibria do exist in many settings of interest. For example, consider symmetric games. Then, either the symmetric equilibrium whose existence is guaranteed by a fixed-point argument is itself efficient, or it is inefficient, and then every player obtains a lower payoff in that equilibrium than in the efficient outcome, so assumption III is satisfied.

3 The main idea

In the main text, we focus on the i.i.d. types. In Appendix, we generalize our results to the case when players' types are Markovian. The main idea of equilibria that we are going to construct is to imitate the AGV mechanism (see d'Aspremont and Gerard-Varet (1979) and Arrow (1979)) using continuation payoffs as transfers. To introduce this idea, suppose for a moment that players are allowed to make monetary transfers to one another at the end of each period.[8]

Theorem 1. *If the stage game satisfies assumptions (A)-(C), and players are allowed to make monetary transfers at the end of each period, then the efficient payoff can be attained in penalty-card equilibria as the discount factor tends to* 1.

This result is not new. For i.i.d. types, it is implicit in Fudenberg et al. (1994). A working paper by Miller (2009) makes this observation explicit, and shows that the result is not limited to i.i.d. types. For Markov types it follows from Athey and Segal's (2013) Proposition 2. Our objective, however, is to prove the result without transfers and in explicit and relatively intuitive strategies, improving in this way on both Fudenberg et al. and Athey and Segal. We provide the proof of Theorem 1 not only for completeness, we will refer to this construction in the following sections.

First, we introduce some auxiliary terms, which will also be used later. Define by

$$s_i^j = E_{\theta_{-i}}(u_j(\theta_j, a(\theta_i, \theta_{-i}))) - E_{\theta_i} E_{\theta_{-i}}(u_j(\theta_j, a(\theta_i, \theta_{-i})))$$

the effect of player i's report on player j's payoff; in particular, $s_i^j > 0$ ($s_i^j \leq 0$) if player i reports a type that gives player j in expectation a payoff higher (no higher) than the ex ante expected payoff. This effect is obviously a function of θ_i, but we will often disregard its argument as it will cause no confusion. Let

$$s_i = \sum_{j \neq i} s_i^j$$

be the effect of player i's report on the total payoff of all other players.

[8] In this case, players need not observe any public randomization device.

To show Theorem 1, consider the following strategies. To simplify notation for any type profile θ, denote by $a(\theta)$ the *efficient action profile*, i.e., an action profile that maximizes the sum of the stage-game payoffs of all players. (This is, the R-efficient action profile $a(\theta_R)$ for $R = \{1, ..., I\}$.)

(A) In every period, players report their types truthfully.

(B) If θ is the reported type profile, players take action profile $a(\theta)$.

(C) Players make transfers. For all $i \neq j \in \{1, ..., I\}$, player j transfers s_i^j to player i.

That is, player i obtains (as a transfer) the difference between the sum of interim and ex ante expected payoffs of other players. Player i's expected payoff from reporting θ_i', given truthful reports of other players, is then

$$E_{\theta_{-i}}(u_i(\theta_i, a(\theta_i', \theta_{-i}))) + \sum_{j \neq i} \left[E_{\theta_{-i}}(u_j(\theta_j, a(\theta_i', \theta_{-i}))) - E_{\theta_i} E_{\theta_{-i}}(u_j(\theta_j, a(\theta_i, \theta_{-i})))\right] -$$

$$- \sum_{j \neq i} E_{\theta_j} \left[E_{\theta_{-j}}(u_i(\theta_i, a(\theta_i, \theta_{-i}))) - E_{\theta_j} E_{\theta_{-j}}(u_i(\theta_i, a(\theta_i, \theta_{-i}))\right]).$$

The first term of this expression is player i's expected interim utility given his actual and reported type, the second term is the expected payment to player i from other players, and the third term is the expected payment of player i to other players. The third term is equal to zero, and the second part of the second term does not depend on player i's report, while the first term and the first part of the second term sum up to

$$\sum_{j=1}^{I} E_{\theta_{-i}}(u_j(\theta_j, a(\theta_i', \theta_{-i}))) \tag{1}$$

Thus, if players other than i report truthfully, player i has an incentive to maximize the sum of the stage-game payoffs, which is attained by reporting her own type truthfully.

(D) An action profile other than $a(\theta)$, for any reported type profile θ, triggers a permanent repetition of the bad stage-game Nash equilibrium. This obviously disciplines the players to taking action profile $a(\theta)$ given any report θ.

That is, the prescribed strategies are incentive compatible, and attain the efficient payoff.

Transfers are not allowed in our setting, they will be dispensed with later, and their role will be played by continuation payoffs. More specifically, a player i's report will affect the probability of i obtaining a penalty card (or of annulling i's existing penalty cards), which in turn increases (decreases) the probability of i going on suspension. Since players on suspension will play actions that maximize the stage-game payoffs of active players, i's suspension can be viewed as a transfer to other players. This transfer can be player j's specific by making the probability of i obtaining a penalty card dependent on player j's report, and different for different players j.

4 Efficient repeated-game strategies

In the present section, we specify penalty-card strategies that attain efficiency. That is, we show that:

Theorem 2. *If players types are i.i.d., and the stage game satisfies assumptions I-III, then the efficient payoff can be approximated in penalty-card equilibria when the discount factor δ approaches* 1.

Let

$$p_i = \Pr\{s_i > 0\} \cdot E_{\theta_i}[s_i | s_i > 0],$$

or equivalently,

$$p_i = -\Pr\{s_i \leq 0\} \cdot E_{\theta_i}[s_i | s_i \leq 0].$$

We first describe the efficient strategies in the case when all players are active.

At the beginning of period 1, a player is randomly selected, each with probability $1/I$, and that player begins the game with one penalty card.[9] There is always exactly one player who holds a positive number of penalty cards. Denote this player i by F. At the end of each period, all penalty cards of player F can be annulled, player F can obtain another penalty card, or the penalty-card structure may stay unchanged. When all penalty cards of player F are annulled, another player j obtains a (first) penalty card; denote this other player by G. It is decided at the end of the period (contingent on the realization of public randomization device) who will be player G; each player other than F has a chance of $1/(I-1)$ of becoming player G.

Denote by n the common limit on the number of penalty cards that an active player can hold, that is, a player who obtains the n-th penalty card goes on suspension. Suppose that player i is player F at the beginning of the current period, and that she holds $k < n$ cards. Denote the three possible penalty-card structures at the end of the period by O_G^1, O_F^{k+1} and O_F^k, respectively.

(A&B) As in Section 3, players report their types truthfully, and if θ is the reported type profile, players take action profile $a(\theta)$.

The penalty-card structure in the following period is determined at the end of the period contingent on the realization of public randomization device, by the following four-component lottery. The first two components describe terms that depend on the players' reports of their types in the current period. The third and fourth components are adjustment term, independent of the reports.

(C1) O_G^1 with probability $\alpha_k^i s_i \mathbf{1}_{\{s_i > 0\}}$, and O_F^{k+1} with probability $-\phi_k^i s_i \mathbf{1}_{\{s_i \leq 0\}}$.

Numbers α_k^i and ϕ_k^i will be specified later. For now, it is important to know that they will converge to 0 as the discount factor converges to 1.

(C2) O_G^1 with probability $-\psi_{i,k}^j s_j \mathbf{1}_{\{s_j \leq 0\}}$, and O_F^{k+1} with probability $s_j \beta_{i,k}^j \mathbf{1}_{\{s_j > 0\}}$.

Numbers $\beta_{i,k}^j$ and $\psi_{i,k}^j$ will be specified later, and will also converge to 0 as the discount factor converges to 1.

[9] To define the strategies in this manner, we need to allow players to observe a realization of public randomization device at the beginning of period 1. However, this is not necessary, as one can also choose player F in a deterministic manner, which introduces some asymmetry; this would be inconvenient for the analysis of players' payoffs and incentives, but the result would still hold true.

(C3) O_F^{k+1} with probability

$$\frac{k}{2(k+1)} - \phi_k^i p_i - \frac{1}{I-1}\sum_{j\neq i}\beta_{i,k}^j p_j,$$

independent of the messages sent (or the actions played) in the current period; O_G^1 with probability

$$\frac{1}{I-1}\left(\frac{1}{2(k+1)} - \alpha_k^i p_i - \psi_{i,k}^j p_j\right),$$

for a given j, independent of the messages and actions in the current period;[10]

(C4) O_F^k with the remaining probability.

(D) As in Section 3, if an action profile other than $a(\theta)$ is observed, for any reported type profile θ, players switch in the following period to playing permanently the bad stage-game Nash equilibrium.

(E) If some player i reaches the limit of n penalty cards, she goes on suspension. This means that for the expected number of M periods players report truthfully their types, and play the R-efficient action profile $a(\theta_R)$, where $R = \{1, ..., I\} - \{i\}$. After the M periods, player i comes back from suspension, which means that one penalty card of player i is annulled.[11]

This last prescription of play applies only under the assumption that no other player goes on suspension during the M periods. We will make this simplifying assumption until Section 6, in which we specify the details of play when a player goes on suspension.

Remark 1 *Note that under these strategies the expected change in the number of cards of player F is* 0*. Player F obtains one more card with probability* $k/2(k+1)$ *and all* k *cards are annulled with probability* $1/2(k+1)$*. We tried several other combinations of transition probabilities in the penalty-card structure, but the proof was always unravelling. If player F was obtaining cards too quickly, we were losing efficiency, as the probability of player F going on suspension was too high. If player F was obtaining cards too slowly, players had insufficient incentives for revealing her type truthfully.*

5 Analysis

5.1 Value functions

It will be convenient to adopt a slightly simpler notation. Namely, let $v^i = v_R^i$ for $R = \{1, ..., I\} - \{i\}$ be the stage game payoff of player i when she is the only player on suspension, and let $w_j^i = w_R^i$ for $R = \{1, ..., I\} - \{j\}$ be the stage game payoff of player i when some other player j is the only player on suspension, and let $w^i = w_{\{1,...,I\}}$ be the stage game payoff of player i when all players are active. Denote by V_k^i the continuation payoff of player i who is currently player F and holds k cards, and by $W_{j,k}^i$ the continuation payoff of player i when j is currently player F and holds k cards. These payoffs are obviously

[10] As numbers α_k^i, ϕ_k^i, $\beta_{i,k}^j$, $\psi_{i,k}^j$ will be small, the formulas in the displays define positive numbers.

[11] It is worth pointing out that truthful reporting is incentive compatible for players on suspension, since their reports have no effect on R-efficient action profiles (by the assumption that each player's payoffs are independent of other players types).

computed assuming that players play the prescribed strategies. We will often call V_k^i and $W_{j,k}^i$ *value functions.* These functions are payoffs at an ex ante stage, when players have not yet learned their current types.

Numbers α_k^i and ϕ_k^i, which will sometimes be called *probabilities of control*, will be defined so that the following equalities are satisfied:

$$\alpha_k^i s_i(1-\varepsilon)\left[\frac{1}{I-1}\sum_{j\neq i} W_{j,1}^i - V_k^i\right] = s_i\varepsilon$$

for $s_i > 0$, and

$$\phi_k^i(-s_i)(1-\varepsilon)[V_{k+1}^i - V_k^i] = s_i\varepsilon$$

for $s_i \leq 0$.

This choice of α_k^i and ϕ_k^i gives player i an incentive to maximize the stage-game payoffs of all players, in the way analogous to transfers from Section 3. Indeed, by (C1) and (C2) of the definition of strategies, player i's report affects her continuation payoff in the "subgame"[12] beginning in the following period through the left-hand sides of the equations. In turn, the right-hand sides are equal to the sum of stage-game payoffs across all players other than i, which together with the effect of player i's report on her stage-game payoff yields the desired incentive to maximize the sum of the stage-game payoffs of all players.

By dividing the two equations by s_i we obtain

$$\alpha_k^i(1-\varepsilon)\left[\frac{1}{I-1}\sum_{j\neq i} W_{j,1}^i - V_k^i\right] = \varepsilon \text{ and } \phi_k^i(1-\varepsilon)[V_k^i - V_{k+1}^i] = \varepsilon. \tag{2}$$

Similarly, numbers $\beta_{i,k}^j$ and $\psi_{i,k}^j$ (which will also be sometimes called probabilities of control) will be defined so that the following equations are satisfied:

$$\frac{1}{I-1}\beta_{j,k}^i s_i(1-\varepsilon)[W_{j,k+1}^i - W_{j,k}^i] = s_i\varepsilon$$

for $s_i > 0$, and

$$\frac{1}{I-1}\psi_{j,k}^i(-s_i)(1-\varepsilon)\left[V_1^i - W_{j,k}^i\right] = s_i\varepsilon$$

for $s_i \leq 0$.

These equations guarantee that players who currently hold no card maximize the total (across all players) stage-game payoff, and by dividing by s_i we obtain

$$\frac{1}{I-1}\beta_{j,k}^i(1-\varepsilon)[W_{j,k+1}^i - W_{j,k}^i] = \varepsilon \text{ and } \frac{1}{I-1}\psi_{j,k}^i(1-\varepsilon)\left[W_{j,k}^i - V_1^i\right] = \varepsilon. \tag{3}$$

Given the prescribed strategies, value V_k^i for $k = 1, ..., n-1$ satisfies the following recursive equation:

$$V_k^i = \varepsilon w^i + (1-\varepsilon)\frac{1}{2}V_k^i + (1-\varepsilon)\frac{k}{2(k+1)}V_{k+1}^i + (1-\varepsilon)\frac{1}{I-1}\frac{1}{2(k+1)}\sum_{j\neq i} W_{j,1}^i.$$

[12] Strictly speaking, the term subgame is incorrect, since private types are never revealed. We will disregrad this subtle issue, which should cause no confusion.

Indeed, player i's current stage game payoff is w^i. By (C1) of the definition of strategies, player i obtains another penalty card with probability $-s_i\phi^i_k$ when $s_i \leq 0$; in expectation, this yields $\phi^i_k p_i$. By (C2) of the definition, player i obtains another penalty card with probability $\beta^j_{i,k}p_j$ when player j, as player G, decides about the penalty-card structure. (Recall that player j learns whether she is player G after reporting the current type.) Together with the chance of obtaining another card described in (C3) of the definition, this yields $k/2(k+1)$. Similarly, we compute the chance of all player i's cards being annulled (in which case player G obtains a penalty card), and the chance that the number of cards of player i stay the same.

Moving the term $(1-\varepsilon)V^i_k/2$ to the left-hand side, then dividing the equation by $1-(1-\varepsilon)/2$, and omitting all terms of order smaller than ε,[13] one can rewrite the recursive equation as:

$$V^i_k = 2\varepsilon w^i + (1-2\varepsilon)\frac{k}{k+1}V^i_{k+1} + (1-2\varepsilon)\frac{1}{I-1}\frac{1}{k+1}\sum_{j\neq i} W^i_{j,1}, \tag{4}$$

and for $k=n$ (also omitting terms smaller than ε):

$$V^i_n = M\varepsilon v^i + (1-M\varepsilon)V^i_{n-1}. \tag{5}$$

We have assumed here that when player i goes on suspension, she will be the only player on suspension for the entire duration of suspension (M periods). When we fully specify the equilibrium strategies in Section 6, this will not be entirely true, that is, several players may be on suspension at the same time. Then, the equation will be satisfied only in approximation, but this will be enough for our purposes. We will return to this issue in Section 6.

Similarly to V^i_k, we obtain value $W^i_{j,k}$ the following recursive equation for $k=1,...,n-1$:

$$W^i_{j,k} = 2\varepsilon w^i + (1-2\varepsilon)\frac{k}{k+1}W^i_{j,k+1} + (1-2\varepsilon)\frac{1}{I-1}\frac{1}{k+1}\sum_{m\neq i,j} W^i_{m,1} + (1-2\varepsilon)\frac{1}{I-1}\frac{1}{k+1}V^i_1,$$

and for $k=n$:

$$W^i_{j,n} = M\varepsilon w^i_j + (1-M\varepsilon)W^i_{j,n-1}.$$

Notice that the recursive equations involve no probability of control. This is important, since it enables us to compute the value functions from the recursive equations, and then define the probabilities of control by equations (2) and (3).

5.2 Efficiency

In this section, we will show that the strategies described in the previous sections achieve efficiency. The calculations will be performed assuming that ε is very small, so one can disregard terms of order lower than ε.

[13] Whem we omit terms of order lower than ε, then dividing by $1-(1-\varepsilon)/2 = 1/2+\varepsilon/2$ is equivalent to multiplying by $2-2\varepsilon$.

From the formulas for V_k^i and V_n^i (that is, (4) and (5)), we can derive the expression for V_{n-1}^i as a function of $W_{j,1}^i$ as follows:

$$V_{n-1}^i = 2\varepsilon w^i + (1-2\varepsilon)\frac{n-1}{n}V_n^i + (1-2\varepsilon)\frac{1}{I-1}\frac{1}{n}\sum_{j\neq i} W_{j,1}^i$$

$$\begin{aligned} &= 2\varepsilon w^i + (1-2\varepsilon)\frac{n-1}{n}[M\varepsilon v^i + (1-M\varepsilon)V_{n-1}^i] + (1-2\varepsilon)\frac{1}{I-1}\frac{1}{n}\sum_{j\neq i} W_{j,1}^i \\ &= 2\varepsilon w^i + \frac{n-1}{n}M\varepsilon v^i + \frac{n-1}{n}[1-(M+2)\varepsilon]V_{n-1}^i + (1-2\varepsilon)\frac{1}{I-1}\frac{1}{n}\sum_{j\neq i} W_{j,1}^i, \end{aligned}$$

which yields

$$V_{n-1}^i = 2n\varepsilon w^i + (n-1)M\varepsilon v^i + \{1-[2n+(n-1)M]\varepsilon\}\frac{1}{I-1}\sum_{j\neq i} W_{j,1}^i.$$

We will now recursively demonstrate that

$$V_k^i = \varepsilon C_k^i + (1-D_k\varepsilon)\frac{1}{I-1}\sum_{j\neq i} W_{j,1}^i \tag{6}$$

for some constants C_k^i and D_k.

Suppose (6) holds for $k+1$. This and (5) yield

$$\begin{aligned} V_k^i &= 2\varepsilon w^i + (1-2\varepsilon)\frac{k}{k+1}\left[\varepsilon C_{k+1}^i + (1-D_{k+1}\varepsilon)\frac{1}{I-1}\sum_{j\neq i} W_{j,1}^i\right] + (1-2\varepsilon)\frac{1}{I-1}\frac{1}{k+1}\sum_{j\neq i} W_{j,1}^i \\ &= \varepsilon\left(2w^i + \frac{k}{k+1}C_{k+1}^i\right) + \left[1-\varepsilon\left(2+\frac{k}{k+1}D_{k+1}\right)\right]\frac{1}{I-1}\sum_{j\neq i} W_{j,1}^i, \end{aligned}$$

this means that (6) holds for k, and

$$C_k^i = 2w^i + \frac{k}{k+1}C_{k+1}^i \text{ and } D_k = 2 + \frac{k}{k+1}D_{k+1}. \tag{7}$$

An analogous argument yields:

$$W_{j,n-1}^i = 2n\varepsilon w^i + (n-1)M\varepsilon w_j^i + \{1-[2n+(n-1)M]\varepsilon\}U_{j,1}^i,$$

where

$$U_{j,1}^i = \frac{1}{I-1}\sum_{m\neq i,j} W_{m,1}^i + \frac{1}{I-1}V_1^i;$$

and

$$W_{j,k}^i = \varepsilon c_{j,k}^i + (1-d_k\varepsilon)U_{j,1}^i, \tag{8}$$

where

$$c_{j,k}^i = 2w^i + \frac{k}{k+1}c_{j,k+1}^i \text{ and } d_k = 2 + \frac{k}{k+1}d_{k+1}.$$

By summing up (8) for $k=1$ across all $j\neq i$, we obtain

$$\frac{1}{I-1}\sum_{j\neq i} W_{j,1}^i = \varepsilon\sum_{j\neq i} c_{j,1}^i + [1-(I-1)d_1\varepsilon]V_1^i.$$

This equation together with (6) for $k = 1$ enables us to compute V_1^i:

$$V_1^i = \frac{C_1^i + \sum_{j \neq i} c_{j,1}^i}{D_1 + (I-1)d_1}, \tag{9}$$

and

$$\frac{1}{I-1} \sum_{j \neq i} W_{j,1}^i = \frac{(1 + D_1 \varepsilon) \sum_{j \neq i} c_{j,1}^i + [1 - (I-1)d_1 \varepsilon] C_1^i}{D_1 + (I-1)d_1}. \tag{10}$$

By recursive computing, we obtain:

$$C_1^i = 2w^i \left(1 + \frac{1}{2} + ... + \frac{1}{n-2}\right) + \frac{2nw^i}{n-1} + Mv^i, \tag{11}$$

$$D_1 = 2\left(1 + \frac{1}{2} + ... + \frac{1}{n-2}\right) + \frac{2n}{n-1} + M, \tag{12}$$

$$c_{j,1}^i = 2w^i \left(1 + \frac{1}{2} + ... + \frac{1}{n-2}\right) + \frac{2nw^i}{n-1} + Mw_j^i,$$

and

$$d_1 = 2\left(1 + \frac{1}{2} + ... + \frac{1}{n-2}\right) + \frac{2n}{n-1} + M.$$

This yields

$$\lim_n V_1^i = w^i \tag{13}$$

and implies that the payoff is efficient, if M goes to infinity at a rate lower than

$$\sum_{m=1}^{n} \frac{1}{m}.$$

Similarly, we obtain that

$$\lim_n \frac{1}{I-1} \sum_{j \neq i} W_{j,1}^i = w^i. \tag{14}$$

5.3 Probabilities of control, incentives

The probabilities of control α_k^i, ϕ_k^i, $\beta_{i,k}^j$ and $\psi_{i,k}^j$ are defined by conditions (2) and (3), where the value functions are determined by the recursive equations from the previous section. Notice that the value functions determined by the recursive equations may not be the actual value functions in the repeated game, but the two value functions will be equal in approximation. It only remains to show that these numbers are positive but small. We will show that this is the case for α_k^i and ϕ_k^i; for $\beta_{i,k}^j$ and $\psi_{i,k}^j$, the argument is analogous.

First, it follows easily from the recursive equations for C_k^i and D_k (7) and the formulas for C_1^i and D_1 (12) and (11) that

$$C_k^i = 2w^i \left(1 + \frac{k}{k+1} + \frac{k}{k+2} + ... + \frac{k}{n-2}\right) + \frac{2knw^i}{n-1} + kMv^i,$$

and

$$D_k = 2\left(1 + \frac{k}{k+1} + \frac{k}{k+2} + ... + \frac{k}{n-2}\right) + \frac{2kn}{n-1} + kM.$$

By (6)

$$V_k^i - \frac{1}{I-1}\sum_{j\neq i} W_{j,1}^i = \varepsilon C_k^i - D_k \varepsilon \frac{1}{I-1}\sum_{j\neq i} W_{j,1}^i,$$

and since by (9) and (10), the difference between $\frac{1}{I-1}\sum_{j\neq i} W_{j,1}^i$ and V_1^i is at most of order ε, which means that, disregarding expressions of order ε^2, one may replace $\frac{1}{I-1}\sum_{j\neq i} W_{j,1}^i$ with V_1^i on the right-hand side. Using (9) and the formulas for C_1^i, D_1, $c_{j,1}^i$ and d_1, we obtain that

$$V_1^i = w^i + \frac{M}{2\left(1+\frac{1}{2}+...+\frac{1}{n-2}\right)+\frac{2n}{n-1}+M}\left(\frac{1}{I}v_i + \frac{1}{I}\sum_{j\neq i} w_j^i - w^i\right). \tag{15}$$

If the expression in parenthesis is negative, then

$$V_k^i - \frac{1}{I-1}\sum_{j\neq i} W_{j,1}^i$$

$$= \varepsilon kM(v^i - w^i) - \varepsilon \frac{[2\left(1+\frac{k}{k+1}+\frac{k}{k+2}+...+\frac{k}{n-2}\right)+\frac{2kn}{n-1}+kM]}{2\left(1+\frac{1}{2}+...+\frac{1}{n-2}\right)+\frac{2n}{n-1}+M} M\left(\frac{1}{I}v^i + \frac{1}{I}\sum_{j\neq i} w_j^i - w^i\right)$$

$$\leq \varepsilon kM(v^i - w^i) - \varepsilon kM\left(\frac{1}{I}v^i + \frac{1}{I}\sum_{j\neq i} w_j^i - w^i\right) = -\varepsilon kM\frac{I-1}{I}\left(\frac{1}{I-1}\sum_{j\neq i} w_j^i - v^i\right),$$

and by assumption II, this is a negative expression at least of order $kM\varepsilon$, which implies that α_k^i is positive but small if M is sufficiently large.

If the expression in parenthesis in (15) is positive, then

$$V_k^i - \frac{1}{I-1}\sum_{j\neq i} W_{j,1}^i < \varepsilon kM(v^i - w^i),$$

and so by assumption I, this is again a negative expression at least of order $kM\varepsilon$.

Similarly,

$$V_k^i - V_{k+1}^i = \varepsilon[C_k^i - C_{k+1}^i] - [D_k - D_{k+1}]\varepsilon\frac{1}{I-1}\sum_{j\neq i} W_{j,1}^i,$$

which by the recursive equations for C_k^i and D_k, and (6) is equal to

$$-\varepsilon\frac{1}{k+1}\left(C_{k+1}^i - D_{k+1}\frac{1}{I-1}\sum_{j\neq i} W_{j,1}^i\right) + 2\varepsilon\left(w^i - \frac{1}{I-1}\sum_{j\neq i} W_{j,1}^i\right)$$

$$= -\frac{1}{k+1}\left(V_{k+1}^i - \frac{1}{I-1}\sum_{j\neq i} W_{j,1}^i\right) + 2\varepsilon\left(w^i - \frac{1}{I-1}\sum_{j\neq i} W_{j,1}^i\right).$$

So, by the previous argument and (14), we also have that ϕ_k^i is positive but small if M is sufficiently large.

6 Play on suspension

To complete the analysis, we need to specify the play when a player is on suspension. First, notice that our analysis up to now is valid if we assume that players go on suspension not for a deterministic number of M periods, but for a random number of periods. More precisely, every period a player on suspension is allowed to come back and become active with probability μ such that

$$M = \sum_{t=1}^{\infty} t\mu(1-\mu)^{t-1} = \frac{1}{\mu}. \tag{16}$$

We define the repeated-game strategies as follows: In the original I-player game, all players are initially active. Once a player i_1 goes on suspension, players start playing an $(I-1)$-player subgame in which they maximize the total payoff of players other than i_1. In each period of the $(I-1)$-player subgame, player i_1 can return from suspension, in which case players resume playing the I-player game, and one card is returned to player i_1 (so i_1 owes $n-1$ cards). What happened in the $(I-1)$-player subgame becomes irrelevant. If player i_1 goes on suspension again, then players start playing the $(I-1)$-player subgame from the beginning, not from the moment they stopped because player i_1 returned from suspension.

It may happen in the $(I-1)$-player subgame that another player i_2 goes on suspension, in which case a $(I-2)$-player subgame is initiated; in this subgame, players maximize the total payoff of active players, that is, all players but i_1 and i_2. In each period either of players i_1 and i_2 may return from suspension. If this is player i_1, then player i_2 returns to the game as well; player i_1 went on suspension earlier (in the I-player game), and when she returns, player i_2 who went on suspension later must return as well. Players resume playing the I-player game, one card is returned to player i_1, and what happened from the time player i_1 went on suspension becomes irrelevant. If this is player i_2 who first returns from suspension, then players resume playing the $(I-1)$-player subgame, one card is returned to player i_2, and only what happened in the $(I-2)$-player subgame becomes irrelevant.

More generally, for any sequence of players $i_1, i_2, ..., i_l$ on suspension, the total payoff of remaining, active players is maximized in the $(I-l)$-player subgame.[14] If player i_k returns from suspension, then all players $i_{k+1}, ..., i_l$ return as well; one card is returned to player i_k and players continue playing the $(I-k+1)$-player subgame. Moreover, what happened after player i_1 went on suspension becomes irrelevant.

From the perspective of the $(I-l)$-player subgame, the probability of interrupting the game, when one of the players $i_1, i_2, ..., i_l$ comes back from suspension is equivalent to additional discounting. This probability, and so the additional discounting vanishes with the discounting in the $(I-l+1)$-player subgame. Thus, the discounting in the $(I-l)$-player subgame vanishes when the discounting in the $(I-l+1)$-player subgame vanishes, although it does so at a lower rate.

Once the joint-penalty phase is reached, the play remains in this phase forever, and players play the bad stage-game equilibrium (see assumption III), even if this phase begins when some players are on suspension.

[14] In particular, if player i is the only active player, then players play the actions that maximize i's payoff, given the reported type of this player. There is no need to keep any penalty-card record until another player returns from suspension.

An inductive argument shows that the expected payoff of every player i in every subgame at the beginning of which the play is in the cooperation phase converges to the efficient payoff w^i. Therefore, players have incentives to maintain the play in cooperation phase.

Finally, we need to justify the recursive formulas for V_n^i and $W_{j,n}^i$. Those formulas were obtained under the assumption that when a player is on suspension, player i obtains for M periods the stage-game payoff of v^i or w_j^i (depending on whether i or $j \neq i$ is on suspension). This, however, is indeed true in approximation. It can be easily proved by induction (with respect to the number of active players) that the R-efficient payoff vector is attained in the subgame with R being the subset of active players, when the discount factor δ tends to 1, and the probability of any player returning from suspension tends to 0.

7 Advantages of penalty-card strategies

7.1 Low number of states

One advantage of penalty-card strategy profiles over other strategies invented for similar purposes in the existing literature, such as review strategies, is that players condition their actions only on a simple statistics of the past play, especially when the number of players is small. Namely, they condition on whether the play is in the cooperation or joint-penalty phase, who are the players on suspension and the order in which they went on suspension, and who currently holds penalty cards and how many of them.

In particular, the penalty-card strategy profiles only minimally, indirectly depend on the space of types and actions. Therefore, penalty-card strategies seem particularly attractive in games with a small number of players, and large numbers of types and actions. In contrast, review strategy profiles typically prescribe different actions for different types, and when the number of both types and actions is large, then that strategy profile requires performing a large number of frequency tests that check whether the action prescribed for each type has been played with roughly "right" frequency.

7.2 Other advantages, and limitations

We believe that penalty-card strategies, or more generally "debt strategies" dominate the review strategies, in the sense that any test used in a review can induce only some specific features of behavior, in specific settings, while debt is a more universal way of providing incentives. For example, tests which induce desired behavior when types are i.i.d. may fail when the types are Markov. In turn, debt strategies allow players to use their private information in the way that is most beneficial for them, and impose only some limits which enforce paying the debt back.

Our results in this paper provide only partial support for this claim; as one will see in Appendix, the construction of efficient strategies depends on the stochastic process which governs the evolution of players' types. However, we show in the companion paper Olszewski and Safronov (2015) that in numerous applications, the same penalty-card strategies approximate efficiency for all Markov chains with transition probability bounded away from 1.

In terms of limitations, the discussion of Section 7.1 suggests that penalty-card strategies seem inferior to review strategies in games with a large number of players and small numbers of types and actions. In addition, the penalty-card strategies used to prove our general results rely heavily, and in a subtle manner, on public randomization. However, public randomization, and even communication is redundant in many specific settings; again, we refer the reader to our companion paper.

8 Markov types

8.1 Signalling and ratcheting effects

In the Markovian case, the construction of equilibria encounters an additional difficulty. Since types are persistent, reports of types have longer-lasting effects. This may result in "signalling and ratcheting effects," namely, reporting some types may affect future reports of other players, giving players additional incentives to misreport.

To be more specific, recall that in the i.i.d. case player i's report affects the penalty card structure through s_i; we defined s_i as the difference between the total expected stage-game payoff of players $j \neq i$ contingent on player i's type being the reported type, and their prior total expected stage-game payoff. By making the probability of obtaining a penalty card (and so player i's continuation payoffs) a function of s_i, we aligned the player's individual incentives with the objective of maximizing the total payoff of all players.

In the Markov case, the expected payoffs, and so the difference, depend on the profile reported in the previous period. So, one might try to define s_i as the same difference as in the i.i.d. case, but contingent on the previous-period types being the reported types, that is,

$$\sum_{j \neq i} \{E_{\theta_{-i}}[u_j(\theta_j, a(\theta_i, \theta_{-i})) \mid \theta_{-i}^{-1}] - E_\theta[u_j(\theta_j, a(\theta_i, \theta_{-i})) \mid \theta^{-1}]\}, \tag{17}$$

where $\theta^{-1} = (\theta_i^{-1}, \theta_{-i}^{-1})$ denotes the type profile reported in the previous period, and $E_{\theta_{-i}}[\cdot \mid \theta_{-i}^{-1}]$ and $E_\theta[\cdot \mid \theta^{-1}]$ mean that the expected values over the distribution of θ_{-i} and θ contingent on the previous-period types being equal to θ_{-i}^{-1} and θ^{-1}.

Consider the reporting incentives of player i for so defined s_i, given truthful reports of all other agents. The impact of player i's report θ_i^{-1} on her continuation payoff (beginning in the current period) would be ε times expression (17). The first term of this expression does not depend on report θ_i^{-1}, it does depend only on report θ_i. The second term, in turn, is determined before observing the current report θ_i, and depends on player i's previous report θ_i^{-1}.

This means that player i's report θ_i^{-1} affects the value of s_i not only in the period it is reported, but also in the following period. In other words, player i has an additional incentive, compared to the i.i.d. case, to report θ_i^{-1}'s that give low values of $E_\theta[u_j(\theta_j, a(\theta_i, \theta_{-i})) \mid \theta^{-1}]$. These are the signalling and ratcheting effects mentioned earlier.

In order to remove these additional incentives, one might try to add the following term to s_i:

$$(1-\varepsilon)\sum_{j\neq i} E_{\theta^{+1}}[u_j(\theta_j^{+1}, a(\theta_i^{+1}, \theta_{-i}^{+1})) \mid \theta_i, \theta_{-i}^{-1}],$$

where $\theta^{+1} = (\theta_i^{+1}, \theta_{-i}^{+1})$ denotes the next-period type profile. This term removes the additional incentives for misreporting θ_i. However, a new problem appears, namely, the expected value of all these new terms, given θ^{-1}, may not be equal to 0. And the nonzero terms make the probability distribution over penalty cards in the following period depend on the currently reported type profile, which would in turn make the analysis of value functions intractable.

One can restore the tractability of analysis by subtracting the expected value of the new terms. However, this operation again creates additional incentives for misreporting. As a result, one keeps including newer terms to the formula for s_i. Fortunately, these newer terms refer to the expectations of what will happen in more remote future given the current report. Due to the convergence of Markov chains to the limiting distributions at an exponential rate, the dependence of these expectations on current report will be vanishing. Thus, we need to include only a finite number of them to remove (almost entirely) the additional incentives, and preserve the tractability of analysis.

It will be essential that the number of these new terms is finite and bounded for all discount factors, since it will make $s_i\varepsilon$ an expression of order $O(\varepsilon)$, while an infinite number of terms would make $s_i\varepsilon$ an expression of order $O(1)$. However, since we remove the additional incentives only almost entirely, we will have to make the additional assumption that the efficient action profiles are unique. (Note that this assumption was redundant in i.i.d. case, as reports did not have any signalling and ratcheting effects.)

8.2 The result in the Markov case

Recall that by the ergodic theorem, the Markov chain on the space of types has a limiting type distribution η. Define $\overline{v}_R^i$ and $\overline{w}_R^i$ as the expected stage-game payoff of player i who is not in R and players take the R-efficient action profile, and the expected stage-game payoff of player i who is a member of R and players take the R-efficient action profile, respectively, and types are distributed according to η. We make the following assumptions that are analogous, to the i.i.d. case:

Assumption I': For any $i = 1, ..., n$ and R such that $i \notin R$,

$$\overline{v}_R^i < \overline{w}_{R\cup\{i\}}^i.$$

Assumption II': For any $i = 1, ..., n$ and R such that $i \in R$, we have that

$$\frac{1}{|R|-1}\sum_{i\neq j\in R} \overline{w}_{R-\{j\}}^i > \overline{v}_{R-\{i\}}^i.$$

Assumption III': The repeated game has an equilibrium in which the payoff of every player i is lower than $\bar{w}_R^i$ for $R = \{1, ..., I\}$.

We elaborate on assumption III' in the following section. In the Markovian case, we need one additional assumption. The necessity of making this assumption follows from our discussion in the previous section.

Assumption IV: For all type profiles $\theta \in \Theta$, and all subsets $R \subset \{1, ..., I\}$ there is a unique action profile $a_R(\theta)$ that maximizes the total payoff of all players in R.

We can now state the counterpart of Theorem 2 for Markov types:

Theorem 3. *If players types are Markovian, and the stage game satisfies assumptions I'-III' and IV, then the efficient payoff can be approximated in penalty-card equilibria when the discount factor δ approaches 1.*

8.3 Bad repeated-game equilibria

In the analysis of the i.i.d. case, we assumed the existence of a bad stage-game equilibrium, and specified the strategy profile in the joint-penalty phase as playing in every period the bad stage-game equilibrium. When types are Markov, a repetition of stage-game equilibrium may not be a repeated-game equilibrium. This problem has been pointed out in several earlier papers (see, for example, Athey and Bagwell (2008) and Escobar and Toikka (2010)).

Actually, the existence of any repeated-game equilibrium in the Markovian case for general stage games follows only from the recent paper by Escobar and Toikka. (For the oligopoly game, the existence was established by Athey and Bagwell.) The existence can also be established in a simpler way by referring to a fixed-point argument. More precisely, the mapping that assigns to every repeated-game strategy profile the set of best-response profiles satisfies the conditions of the extension of Kakutani's fixed-point theorem to the Hilbert cube.

However, the existence does not yet guarantee that assumption III' is satisfied. Therefore, it must be assumed that there exist equilibria in which every player obtains a lower payoff than in the efficient outcome. Assumption III' is not too restrictive, though. It is relatively easy to construct explicitly some "bad" repeated-game equilibria in many concrete settings (such as the repeated version of Spulber's oligopoly game). In addition, if one is interested in symmetric games, then Theorem 3 delivers efficient strategies by an argument analogous to that from the second last paragraph of Section 2.3.

9 Appendix

The purpose of this appendix is to prove Theorem 3.

9.1 Efficient strategies

The strategies will be similar to those used in the i.i.d. case. As before, we will first describe the strategies in the case when all players are active. At the beginning of period 1, player F is randomly selected, each with probability $1/I$, and that player begins the game with a penalty card. At the beginning of other periods, F is the player who currently holds a positive number of penalty cards. (As before, there will always be only one player holding penalty cards.) Player G is selected randomly from the $I-1$ players other than F, each of them with probability $1/(I-1)$; this selection takes place at the end of each period, contingent on the realization of public randomization device. Suppose that player i is currently player F and holds $k<n$ cards. As before, denote the three possible penalty-card structures at the end of the period by O_G^1, O_F^{k+1} and O_F^k.

(A&B) All players report their types truthfully, and if θ is the reported type profile, players take action profile $a(\theta)$.

The penalty-card structure in the following period is determined at the end of the period contingent on the realization of public randomization device, by the following four-component lottery:

(C1) O_G^1 with probability $\alpha_k^i s_i \mathbf{1}_{\{s_k^i(\theta^{-1},\theta_i)>0\}}$, and O_F^{k+1} with probability $-\phi_k^i s_i \mathbf{1}_{\{s_k^i(\theta^{-1},\theta_i)\leq 0\}}$.

We will define α_k^i, ϕ_k^i and $s_k^i(\theta^{-1},\theta_i)$ later.

(C2) O_G^1 with probability $-\psi_{i,k}^j s_j \mathbf{1}_{\{s_{i,k}^j(\theta^{-1},\theta_j)\leq 0\}}$, and O_F^{k+1} with probability $s_j \beta_{i,k}^j \mathbf{1}_{\{s_{i,k}^j(\theta^{-1},\theta_j)>0\}}$.

We will define $\beta_{i,k}^j$, $\psi_{i,k}^j$ and $s_{i,k}^j(\theta^{-1},\theta_j)$ later.

(C3) O_F^{k+1} with probability

$$\frac{k}{2(k+1)} - \phi_k^i p_k^i(\theta^{-1}) - \beta_{i,k}^j \frac{1}{I-1}\sum_{j\neq i} p_{i,k}^j(\theta^{-1}),$$

independent of the actions played (or messages sent) in the current period; O_G^1 with probability

$$\frac{1}{I-1}\left(\frac{1}{2(k+1)} - \alpha_k^i p_k^i(\theta^{-1}) - \psi_{i,k}^j p_{i,k}^j(\theta^{-1})\right),$$

for a given j, independent of the actions played in the current period;[15]

(C4) O_F^k with the remaining probability.

Again, $p_k^i(\theta^{-1})$ and $p_{j,k}^i(\theta^{-1})$ will be defined later.

(D) If an action profile other than $a(\theta)$ is observed, for any reported type profile θ, players switch in the following period to playing permanently the bad repeated-game equilibrium.

(E) If player i reaches the limit of n penalty cards, she goes on suspension. This means that for the expected number of M periods players report truthfully their types, and play the R-efficient action profile $a(\theta_R)$, where $R=\{1,...,I\}-\{i\}$. When player i comes back from suspension, one penalty card of player i is annulled.

---

[15] As numbers α_k^i, ϕ_k^i, $\beta_{i,k}^j$, $\psi_{i,k}^j$ will be small, the formulas in the displays define positive numbers.

Part (E) applies only under the assumption that no other player goes on suspension during the suspension of player i. For now, we will make this simplifying assumption, postponing for later the details of play when a player goes on suspension. It will be important that M and n diverge to infinity at the rates such that

$$\sum_{m=1}^{n} \frac{1}{m} \approx M^{3/2}.$$

9.2 Missing definitions

We define the probabilities of control α_k^i, ϕ_k^i, $\beta_{j,k}^i$, and $\psi_{j,k}^i$ by equations (2) and (3), where value functions V_k^i and $W_{j,k}^i$ are as in the i.i.d. case with the limit ergodic distribution η being the probability distribution over types. Notice that these probabilities of control are independent of any reports of types. By the analysis of the i.i.d. case, the probabilities of control satisfy the following condition:

$$\alpha_k^i = \frac{A_k^i}{M(k+1)},\ \beta_{j,k}^i = \frac{B_{j,k}^i}{M},\ \phi_k^i = \frac{\Phi_k^i}{M},\ \psi_{j,k}^i = \frac{\Psi_{j,k}^i}{M(k+1)}, \tag{18}$$

where A_k^i, $B_{j,k}^i$, Φ_k^i and $\Psi_{j,k}^i \geq 0$ are bounded by a constant which does not depend on M and n.

Next, we will define $s_k^i(\theta^{-1}, \theta_i)$ and $s_{i,k}^j(\theta^{-1}, \theta_j)$. Suppose first that player i is currently player F, and has k penalty cards. Let

$$B_{k,T}^i(\theta_{-i}^{-1}, \theta_i) = \sum_{t=0}^{T} \sum_{j \neq i} (1-\varepsilon)^t E[u_j^{+t} \mid \theta_i, \theta_{-i}^{-1}].$$

where the expression $E[u_j^{+t} \mid \theta_i, \theta_{-i}^{-1}]$ stands for the expectation of the stage-game payoff u_j of player j in t periods from now, given the current type θ_i of player i and the previous types θ_{-i}^{-1} of players other than i. This expression represents the impact of player i's current report on the payoffs of all other players in the following T periods, assuming that all players play the prescribed strategies. Let

$$s_{k,T}^i(\theta^{-1}, \theta_i) = B_{k,T}^i(\theta_{-i}^{-1}, \theta_i) - E_{\theta_i}[B_{k,T}^i(\theta_{-i}^{-1}, \theta_i) \mid \theta_i^{-1}].$$

Let $s_k^i(\theta^{-1}, \theta_i) = s_{k,T}^i(\theta^{-1}, \theta_i)$, where T will be defined in a moment. We define $s_{j,k}^i(\theta^{-1}, \theta_i)$ in a similar manner, as the impact of player i's current report when player j is currently player F and currently has k cards. Notice that $s_k^i(\theta^{-1}, \theta_i)$ may differ from $s_{j,k}^i(\theta^{-1}, \theta_i)$, for example, because the chance that player i will be on suspension in t periods ahead depends on who (i or j) currently holds the k cards.

In order to define T, observe that the impact of θ_i on u_j^{+t} vanishes in the remote future. More precisely, we have that

Claim 1. *For any* $\Delta > 0$, any *player* i, *and any types* θ_i', θ_i'' and θ_{-i}, *there exists a number* T *such that for any* $t > T$ we have

$$\left| \sum_{j \neq i} E[u_j^{+t} \mid \theta_i', \theta_{-i}^{-1}] - \sum_{j \neq i} E[u_j^{+t} \mid \theta_i'', \theta_{-i}^{-1}] \right| < \Delta.$$

If players were never going on suspension, this claim would follow directly from the fact that for any two current type profiles the probability that the types profiles will coincide t periods from now tends to 1 at an exponential rate when t grows large. Since players may go on suspension, it may happen that for some type profile $(\theta_i', \theta_{-i}^{-1})$, the probability that player i will be on suspension t periods from now is higher, while for the other type profile, $(\theta_i'', \theta_{-i}^{-1})$, the probability that player i will be on suspension t periods from now is lower. So, the payoffs u_j^{+t} of players $j \neq i$ may be different. However, the probability that player's report will affect going on suspension is of order $O(1/M)$, and the rate of convergence of type profiles over time is independent of M. Therefore, we can assume that M is sufficiently large so that the possibility of a player going on suspension affects the payoffs of other players only marginally.

We can now define T as the number from Claim 1 for any Δ lower than the difference $\sum_{i \in R} [u_i(\theta_i, a(\theta_R)) - u_i(\theta_i, a)]$ for all profiles θ, subsets of players R, and actions $a \neq a(\theta_R)$. By assumption IV this difference in positive.

Finally, let

$$p_k^i(\theta^{-1}) = \Pr\{s_k^i(\theta^{-1}, \theta_i) > 0\} \cdot E_{\theta_i}[s_k^i(\theta^{-1}, \theta_i) | s_k^i(\theta^{-1}, \theta_i) > 0];$$

similarly, let $p_{j,k}^i(\theta^{-1}) = \Pr\{s_{j,k}^i(\theta^{-1}, \theta_i) > 0\} \cdot E_{\theta_i}[s_{j,k}^i(\theta^{-1}, \theta_i) | s_{j,k}^i(\theta^{-1}, \theta_i) > 0]$.

9.3 Efficiency

Given the penalty-card strategies, we define $V_{k,\theta}^i$ as the continuation payoff of player i, at the beginning of a period when she does not know yet the current type, she holds k penalty cards, and the type profile in the previous period was θ. Similarly, we define $W_{j,k,\theta}^i$ as the continuation payoff of player i when player j holds k penalty cards. These values are defined for the true type profile θ, or assuming that players reported their types truthfully. When a player returns from suspension, these values are different, and are denoted by $\hat{V}_{n-1,\theta}^i, \hat{W}_{j,n-1,\theta}^i$; in this case θ denotes the type profile in the last period before suspension.

The recursive equations for the value functions in their exact form are long and complicated, but we will now show that they coincide with the recursive equations in the i.i.d. case up to a factor that vanishes with the discount rate.

The value function $V_{k,\theta}^i$, for $k < n-1$, is equal to the expectation over the type profile θ' in the current period of the sum of the stage-game payoff $\varepsilon w_{\theta'}^i$ and the continuation payoff. The continuation payoff is in turn the sum of the following expressions:

$$(1-\varepsilon)\frac{1}{I-1}\sum_{j \neq i}\Big[\alpha_k^i s_k^i(\theta, \theta_i')\chi\{s_k^i(\theta, \theta_i') > 0\} - \psi_{i,k}^j s_{i,k}^j(\theta, \theta_j')\chi\{s_{i,k}^j(\theta, \theta_j') \leq 0\}$$

$$+\left(\frac{1}{2(k+1)} - \alpha_k^i p_k^i(\theta) - \psi_{i,k}^j p_{i,k}^j(\theta)\right)\Big] W_{j,1,\theta'}^i,$$

where $\chi\{\} \in \{0,1\}$ is the indicator of whether the condition in $\{\}$ is satisfied,

$$(1-\varepsilon)\left[-\phi_k^i s_k^i(\theta,\theta_i')\chi\{s_k^i(\theta,\theta_i')\le 0\} + \frac{1}{I-1}\sum_{j\neq i}\beta_{i,k}^j s_{i,k}^j(\theta,\theta_j')\chi\{s_{i,k}^j(\theta,\theta_j')>0\}\right.$$
$$\left.+\left(\frac{k}{2(k+1)}-\phi_k^i p_k^i(\theta)-\frac{1}{I-1}\sum_{j\neq i}\beta_{i,k}^j p_{i,k}^j(\theta)\right)\right]V_{k+1,\theta'}^i,$$

and

$$(1-\varepsilon)\left[\frac{1}{2}+\alpha_k^i\left(p_k^i(\theta)-s_k^i(\theta,\theta_i')\chi\{s_k^i(\theta,\theta_i')>0\}\right)\right.$$
$$+\frac{1}{I-1}\sum_{j\neq i}\psi_{i,k}^j(p_{i,k}^j(\theta)+s_{i,k}^j(\theta,\theta_j')\chi\{s_{i,k}^j(\theta,\theta_j')\le 0\})+\phi_k^i(p_k^i(\theta)+s_k^i(\theta,\theta_i')\chi\{s_k^i(\theta,\theta_i')\le 0\})$$
$$\left.+\frac{1}{I-1}\sum_{j\neq i}\beta_{i,k}^j(p_{i,k}^j(\theta)-s_{i,k}^j(\theta,\theta_j')\chi\{s_{i,k}^j(\theta,\theta_j')>0\})\right]V_{k,\theta'}^i.$$

The computation of continuation payoffs follows directly from the prescribed strategies. For example, the first expression refers to the situation that a player other than i will hold a penalty card in the following period. This situation happens: (a) when player F's report determines the penalty-card structure and $s_k^i(\theta,\theta_i')>0$; in this case, it happens with probability $\alpha_k^i s_k^i(\theta,\theta_i')$ (see (C1) of the definition of the prescribed strategies); (b) when player G's report determines the penalty-card structure and $s_{i,k}^j(\theta,\theta_j')\le 0$; in this case, it happens with probability $-\psi_{i,k}^j s_{i,k}^j(\theta,\theta_j')$ (see (C2) of the definition of the prescribed strategies); (c) independently of players' reports with probability given in (C3) of the definition of the prescribed strategies.

Thus, the recursive equation for $V_{k,\theta}^i$, $k<n-1$, has the form

$$V_{k,\theta}^i=\sum_{\theta'}\eta_{\theta,\theta'}\left\{\varepsilon w_{\theta'}^i+(1-\varepsilon)\frac{1}{I-1}\sum_{j\neq i}\frac{1}{2(k+1)}(1+J_{j,k,\theta,\theta'}^i)W_{j,1,\theta'}^i\right.$$
$$\left.+(1-\varepsilon)\frac{k}{2(k+1)}(1+I_{k,\theta,\theta'}^i)V_{k+1,\theta'}^i+(1-\varepsilon)\frac{1}{2}\left(1-\frac{1}{I-1}\sum_{j\neq i}\frac{1}{k+1}J_{j,k,\theta,\theta'}^i-\frac{k}{k+1}I_{k,\theta,\theta'}^i\right)V_{k,\theta'}^i\right\},$$

where $J_{j,k,\theta,\theta'}^i, I_{k,\theta,\theta'}^i$ are terms of order $O(1/M)$ by (18), and because $s_k^i(\theta,\theta_i')$ is bounded across all values of ε. That is, $V_{k,\theta}^i$ is a sum of a term of order ε, and a weighted average of $W_{j,1,\theta'}^i, V_{k+1,\theta'}^i, V_{k,\theta'}^i$. In addition, we have that

$$\sum_{\theta'}\eta_{\theta,\theta'}J_{j,k,\theta,\theta'}^i=\sum_{\theta'}\eta_{\theta,\theta'}I_{k,\theta,\theta'}^i=0,$$

so the ex ante probability that $W_{j,1,\theta'}^i$, for some θ', will be the following period continuation payoff is $\frac{1}{I-1}\frac{1}{2(k+1)}$, and the ex ante probability that $V_{k+1,\theta'}^i$, for some θ', will be the following period continuation payoff is $\frac{k}{2(k+1)}$, and with the remaining probability $V_{k,\theta'}^i$, for some θ', will be the following period continuation payoff.

Let $\bar{V}^i_k = \sum_\theta \eta(\theta) V^i_{k,\theta}$ and $\bar{W}^i_{j,k} = \sum_\theta \eta(\theta) W^i_{j,k,\theta}$. Then, by stability of η,[16] we have

$$\sum_\theta \eta(\theta) \sum_{\theta'} \eta_{\theta,\theta'} \varepsilon w^i_{\theta'} = \sum_{\theta'} \eta(\theta') \varepsilon w^i_{\theta'} = \varepsilon \overline{w}^i,$$

$$\sum_\theta \eta(\theta) \sum_{\theta'} \eta_{\theta,\theta'} \frac{1}{2(k+1)} \frac{1}{I-1} \sum_{j \neq i} W^i_{j,1,\theta'} = \frac{1}{2(k+1)} \frac{1}{I-1} \sum_{j \neq i} \bar{W}^i_{j,1},$$

and

$$\sum_\theta \eta(\theta) \sum_{\theta'} \eta_{\theta,\theta'} \frac{k}{2(k+1)} V^i_{k+1,\theta'} = \frac{k}{2(k+1)} \bar{V}^i_{k+1}.$$

This yields, by summing expressions $V^i_{k,\theta}$ with weights $\eta_{\theta,\theta'}$,

$$\begin{aligned}
\bar{V}^i_k = {} & \varepsilon \overline{w}^i + (1-\varepsilon) \frac{1}{2(k+1)} \frac{1}{I-1} \sum_{j \neq i} \bar{W}^i_{j,1} + (1-\varepsilon) \frac{k}{2(k+1)} \bar{V}^i_{k+1} + (1-\varepsilon) \frac{1}{2} \bar{V}^i_k \\
& + \sum_\theta \eta(\theta) \sum_{\theta'} \eta_{\theta,\theta'} \Bigg\{ (1-\varepsilon) \frac{1}{2(k+1)} \frac{1}{I-1} \sum_{j \neq i} J^i_{j,k,\theta,\theta'} W^i_{j,1,\theta'} \\
& + (1-\varepsilon) \frac{k}{2(k+1)} I^i_{k,\theta,\theta'} V^i_{k+1,\theta'} - (1-\varepsilon) \frac{1}{2} \left(\frac{1}{k+1} \frac{1}{I-1} \sum_{j \neq i} J^i_{j,k,\theta,\theta'} + \frac{k}{k+1} I^i_{k,\theta,\theta'} \right) V^i_{k,\theta'} \Bigg\}.
\end{aligned}$$

Since $\sum_{\theta'} \eta_{\theta,\theta'} J^i_{j,k,\theta,\theta'} = 0$, the term $\sum_\theta \eta(\theta) \sum_{\theta'} \eta_{\theta,\theta'} \frac{1}{2(k+1)} \frac{1}{I-1} \sum_{j \neq i} J^i_{j,k,\theta,\theta'} W^i_{j,1,\theta'}$ is a weighted sum of differences $W^i_{j,1,\theta'} - W^i_{j,1,\theta''}$. The following claim shows that these differences are of order $O(\varepsilon)$, and since $J^i_{j,k,\theta,\theta'}$ is of order $O(1/M)$, we have that $\sum_\theta \eta(\theta) \sum_{\theta'} \eta_{\theta,\theta'} \frac{1}{2(k+1)} \frac{1}{I-1} \sum_{j \neq i} J^i_{j,k,\theta,\theta'} W^i_{j,1,\theta'}$ is of order $O(\varepsilon/M)$.

Claim 2. *For any players i and j, number of cards k, and type profiles θ and θ' there is constant $C > 0$, independent of ε, M, n such that $|V^i_{k,\theta} - V^i_{k,\theta'}| < C\varepsilon$ and $|W^i_{j,k,\theta} - W^i_{j,k,\theta'}| < C\varepsilon$. Analogous estimates hold for $\hat{V}^i_{n-1,\theta}$ and $\hat{W}^i_{j,n-1,\theta}$.*[17]

By analogous arguments applied to other terms of the formula for $\bar{V}^i_k$, we obtain that

$$\bar{V}^i_k = \varepsilon \overline{w}^i + (1-\varepsilon) \frac{1}{I-1} \frac{1}{2(k+1)} \sum_{j \neq i} \bar{W}^i_{j,1} + (1-\varepsilon) \frac{k}{2(k+1)} \bar{V}^i_{k+1} + (1-\varepsilon) \frac{1}{2} \bar{V}^i_k + O\left(\varepsilon/M\right).$$

If we disregard the terms of order lower than ε, we can transform this formula into

[16] By stability we mean that

$$\eta(\theta') = \sum_\theta \eta(\theta) \eta_{\theta,\theta'}$$

for all θ'. By the ergodic theorem, the limit distribution of any Markov chain has this property.

[17] This claim follows from two facts: (a) for any two initial type profiles, the probability that the types profiles will coincide t periods from now tends to 1 at an exponential rate, independent of the discount factor, when t grows large; and (b) given the prescribed strategies, for any current card structure the distribution over card structures in the following periods is independent of the previous type profiles θ and θ'.

$$\bar{V}_k^i = 2\varepsilon \overline{w}_k^i + (1-2\varepsilon)\frac{1}{I-1}\frac{1}{k+1}\sum_{j\neq i}\bar{W}_{j,1}^i + (1-2\varepsilon)\frac{k}{k+1}\bar{V}_{k+1}^i,$$

where $\overline{w}_k^i$ differs from $\overline{w}^i$ by a term of order $O(1/M)$.

It appears that this formula for $\bar{V}_k^i$, differs from the formula for V_k^i in the i.i.d. case only by replacing $\overline{w}^i$ with $\overline{w}_k^i$. One can perform similar calculations for $\bar{W}_{j,k}^i$, and then repeat the reasoning from the i.i.d. case to obtain that

$$\bar{V}_1^i = \overline{w}^i + O(1/M^{1/2}) \text{ and } \frac{1}{I-1}\sum_{j\neq i}\bar{W}_{j,1}^i = \overline{w}^i + O(1/M^{1/2}).$$

This, together with Claim 2 implies that our penalty-card strategies attain, as the discount factor tends to 1 (and n and M diverge to infinity), the efficient payoffs.

9.4 Incentives

Observe first that $\bar{V}_k^i$ and $\bar{W}_{j,k}^i$ can be determined by the same system of equations as V_k^i and $W_{j,k}^i$ from the i.i.d. case, except $\overline{w}^i$ replaced with $\overline{w}_k^i$, and the differences between $\overline{w}^i$ and $\overline{w}_k^i$ are of order $O(1/M)$. Therefore, by the same calculations as in the i.i.d. case (see formulas (2) and (3)), we obtain that

$$\alpha_k^i(1-\varepsilon)\left[\frac{1}{I-1}\sum_{j\neq i}\bar{W}_{j,1}^i - \bar{V}_k^i\right] = \varepsilon(1+O(1/M));$$

$$\phi_k^i(1-\varepsilon)[\bar{V}_k^i - \bar{V}_{k+1}^i] = \varepsilon(1+O(1/M));$$

$$\frac{1}{I-1}\beta_{j,k}^i(1-\varepsilon)[\bar{W}_{j,k+1}^i - \bar{W}_{j,k}^i] = \varepsilon(1+O(1/M));$$

$$\frac{1}{I-1}\psi_{j,k}^i(1-\varepsilon)\left[\bar{W}_{j,k}^i - \bar{V}_1^i\right] = \varepsilon(1+O(1/M)).$$

Denote the current actual type profile by θ'. Assume for now that no player is going on suspension within T periods. We will argue in the next section that this simplifying assumption is inessential. By inspection of the formula for $V_{k,\theta}^i$, one can see that player i's current report $\hat{\theta}_i'$ affects: the current payoff $w_{(\theta'_{-i},\hat{\theta}'_i)}^i$, the value of $s_k^i(\theta,\hat{\theta}_i')$, and continuation payoffs $W_{j,1,(\theta'_{-i},\hat{\theta}'_i)}^i$, $V_{k+1,(\theta'_{-i},\hat{\theta}'_i)}^i$ and $V_{k,(\theta'_{-i},\hat{\theta}'_i)}^i$. More specifically, $s_k^i(\theta,\hat{\theta}_i')$ is affected through its first component $B_{k,T}^i(\theta_{-i},\hat{\theta}_i')$, and the continuation payoffs are affected through the value of $E_{\theta_i^{+1}}[B_{k,T}^i(\theta'_{-i},\theta_i^{+1}) \mid \hat{\theta}_i']$.

We will first estimate the effect of player i's report $\hat{\theta}_i'$ on $W_{j,1,(\theta'_{-i},\hat{\theta}'_i)}^i$. By referring to the one-stage deviation principle, we will assume that player i will report truthfully in the future, so the distribution of her future reports will be determined by her true type θ_i', rather than the reported type $\hat{\theta}_i'$.

If the future report θ_i^{+1} is such that $B^i_{j,1,T}(\theta'_{-i},\theta_i^{+1}) > E_{\theta_i^{+1}}[B^i_{j,1,T}(\theta'_{-i},\theta_i^{+1}) \mid \hat{\theta}'_i]$, then $W^i_{j,1,(\theta'_{-i},\hat{\theta}'_i)}$ depends on player i's report $\hat{\theta}'_i$ through

$$(B^i_{j,1,T}(\theta'_{-i},\theta_i^{+1}) - E_{\theta_i^{+1}}[B^i_{j,1,T}(\theta'_{-i},\theta_i^{+1}) \mid \hat{\theta}'_i])\frac{1}{I-1}\beta^i_{j,1}[W^i_{j,2,\theta^{+1}} - W^i_{j,1,\theta^{+1}}]$$

$$= (B^i_{j,1,T}(\theta'_{-i},\theta_i^{+1}) - E_{\theta_i^{+1}}[B^i_{j,1,T}(\theta'_{-i},\theta_i^{+1}) \mid \hat{\theta}'_i])\frac{1}{I-1}\beta^i_{j,1}[\bar{W}^i_{j,2} - \bar{W}^i_{j,1} + O(\varepsilon)]$$

$$= (B^i_{j,1,T}(\theta'_{-i},\theta_i^{+1}) - E_{\theta_i^{+1}}[B^i_{j,1,T}(\theta'_{-i},\theta_i^{+1}) \mid \hat{\theta}'_i])(\varepsilon + O(\varepsilon/M)).$$

The first equality follows from the fact that $W^i_{j,1,\theta^{+1}} = \bar{W}^i_{j,1} + O(\varepsilon)$, which in turn follows from Claim 2, and the second equality follows from the observation made at the beginning of this section (the third display) and the fact that $\beta^i_{j,1} = O(1/M)$. Similarly, if $B^i_{j,1,T}(\theta'_{-i},\theta_i^{+1}) \leq E_{\theta_i^{+1}}[B^i_{j,1,T}(\theta'_{-i},\theta_i^{+1}) \mid \hat{\theta}'_i]$, then $W^i_{j,1,(\theta'_{-i},\hat{\theta}'_i)}$ depends on player i's report $\hat{\theta}'_i$ through

$$(B^i_{j,1,T}(\theta'_{-i},\theta_i^{+1}) - E_{\theta_i^{+1}}[B^i_{j,1,T}(\theta'_{-i},\theta_i^{+1}) \mid \hat{\theta}'_i])\frac{1}{I-1}\psi^i_{j,1}[W^i_{j,1,\theta^{+1}} - V^i_{1,\theta^{+1}}]$$

$$= (B^i_{j,1,T}(\theta'_{-i},\theta_i^{+1}) - E_{\theta_i^{+1}}[B^i_{j,1,T}(\theta'_{-i},\theta_i^{+1}) \mid \hat{\theta}'_i])(\varepsilon + O(\varepsilon/M)).$$

The effects of player i's report $\hat{\theta}'_i$ on $V^i_{k+1,(\theta'_{-i},\hat{\theta}'_i)}$ and $V^i_{k,(\theta'_{-i},\hat{\theta}'_i)}$ take the same form. (Recall that we assumed that no player goes on suspension within T periods.) The overall effect of report $\hat{\theta}'_i$ on player i's value function through continuation payoffs must be adjusted by factor $(1-\varepsilon)$, and considered in expectation contingent on θ_{-i}. This yields

$$-(1-\varepsilon)E_{\theta'_{-i}}E_{\theta_i^{+1}}[B^i_{j,1,T}(\theta'_{-i},\theta_i^{+1}) \mid \hat{\theta}'_i, \theta_{-i}](\varepsilon + O(\varepsilon/M)).$$

The effect of player i's report $\hat{\theta}'_i$ through terms $s^i_k(\theta,\hat{\theta}'_i)I\{s^i_k(\theta,\hat{\theta}'_i) > 0\}$ and $s^i_k(\theta,\theta'_i)I\{s^i_k(\theta,\hat{\theta}'_i) \leq 0\}$ turns out to be $B^i_{k,T}(\theta_{-i},\hat{\theta}'_i)(\varepsilon + O(\varepsilon/M))$. Therefore, the total effect of player i's report $\hat{\theta}'_i$, disregarding terms of order lower than ε, is

$$B^i_{k,T}(\theta_{-i},\hat{\theta}'_i)(\varepsilon + O(\varepsilon/M)) - (1-\varepsilon)E_{\theta'_{-i}}E_{\theta_i^{+1}}[B^i_{j,1,T}(\theta'_{-i},\theta_i^{+1}) \mid \hat{\theta}'_i, \theta_{-i}](\varepsilon + O(\varepsilon/M))$$

$$= \varepsilon[B^i_{k,T}(\theta_{-i},\hat{\theta}'_i) - (1-\varepsilon)E_{\theta'_{-i}}E_{\theta_i^{+1}}[B^i_{j,1,T}(\theta'_{-i},\theta_i^{+1}) \mid \hat{\theta}'_i, \theta_{-i}]] + O(\varepsilon/M).$$

The term $O(\varepsilon/M)$ does not affect incentives. Recalling the definition of $B^i_{k,T}(\theta_{-i},\hat{\theta}'_i)$, we obtain that

$$B^i_{k,T}(\theta_{-i},\hat{\theta}'_i) - (1-\varepsilon)E_{\theta'_{-i}}E_{\theta_i^{+1}}[B^i_{j,1,T}(\theta'_{-i},\theta_i^{+1}) \mid \hat{\theta}'_i, \theta_{-i}]$$

$$= \sum_{t=0}^{T}\sum_{j\neq i}(1-\varepsilon)^t E[u_j^{+t} \mid \hat{\theta}'_i, \theta_{-i}] - (1-\varepsilon)\sum_{t=0}^{T}\sum_{j\neq i}(1-\varepsilon)^t E[u_j^{+t+1} \mid \hat{\theta}'_i, \theta_{-i}]$$

$$= \sum_{j\neq i} E[u_j \mid \hat{\theta}'_i, \theta_{-i}] - \sum_{j\neq i}(1-\varepsilon)^{T+1}E[u_j^{+T+1} \mid \hat{\theta}'_i, \theta_{-i}].$$

The first of the two sums is equal to

$$\sum_{j\neq i} E_{\theta_{-i}}[u_j(\theta_j, a(\hat{\theta}_i', \theta_{-i})) \mid \hat{\theta}_i', \theta_{-i}],$$

and together with the effect of player i's report on her current payoff provides the player incentives to maximize the total payoff, while the second sum depends on $\hat{\theta}_i'$ by a value lower than Δ, and therefore is inessential for player i's incentives.

9.5 Play on suspension.

The strategy profile when some players are on suspension is similar to that in the i.i.d. case. The ordering in which players go on suspension is recorded, and the return of a player from suspension means that all players who went on suspension after her also become active. When a return from suspension interrupts a "subgame", it is not continued in the future; that is, if the same set of players happens to be active, they play the subgame from the very beginning (with a random player having one penalty card). However, there are some issues, specific for the Markovian case, that we will now shortly discuss:

(i) When some players are on suspension, values of $B^i_{k,T}(\theta^{-1}_{-i}, \theta_i)$ and $s^i_k(\theta^{-1}, \theta_i)$ for all active players i should include only the payoffs of other active players j, and should not include the payoffs after a return from suspension interrupts the subgame with this particular set of active players.

(ii) Recall that $s^i_{n-1}(\theta^{-1}, \theta_i) = B^i_{n-1,T}(\theta^{-1}_{-i}, \theta_i) - E_{\theta_i}[B^i_{n-1,T}(\theta^{-1}_{-i}, \theta_i) \mid \theta^{-1}_i]$. It is important that when $s^i_{n-1}(\theta^{-1}, \theta_i)$ is computed for the period following a return from suspension, θ^{-1}_i in its second term $E_{\theta_i}[B^i_{n-1,T}(\theta^{-1}_{-i}, \theta_i) \mid \theta^{-1}_i]$ stands for the type of player i in the last period before the suspension. If it were the type in the last period of suspension, player i might have incentives to misreport her type while being on suspension in order to affect $E_{\theta_i}[B^i_{n-1,T}(\theta^{-1}_{-i}, \theta_i) \mid \theta^{-1}_i]$ in the case she becomes active next period. A similar comment applies to $s^j_{i,n-1}(\theta^{-1}, \theta_j)$.

(iii) We estimated the impact of player i's report $\hat{\theta}_i'$ on $W^i_{j,1,(\theta'_{-i},\hat{\theta}'_i)}$, $V^i_{k+1,(\theta'_{-i},\hat{\theta}'_i)}$ and $V^i_{k,(\theta'_{-i},\hat{\theta}'_i)}$ under the assumption that player i will not go on suspension. However, if k is close to the limit number of cards, the impact of the report of player i may be slightly different for three value functions above, because of the possibility of going on suspension; even more, a player's current report affects probability of going on suspension. However, when a player is on suspension, his report does not matter for the value function, and the probability of going on suspension is affected by a player's report only to the order of $O(1/M)$, and therefore both differences are inessential for incentives.

Similarly, some players may return from suspension within T periods, and this will change the impact of player i's current report. However, since M can be chosen large compared to T, the chance of a return from suspension within the interval of length T is again of the order of $O(1/M)$, and therefore is inessential for incentives.

10 References

Abdulkadirouglu, A., and K. Bagwell (2012). "The Optimal Chips Mechanism in a Model of Favors," mimeo.

Abreu, D., D. G. Pearce and E. Stacchetti (1990). "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring," *Econometrica*, 58(5), 1041-1063.

Aoyagi, M. (2003). "Bid Rotation and Collusion in Repeated Auctions," *Journal of Economic Theory*,112, 79-106.

Aoyagi, M. (2007), "Efficient Collusion in Repeated Auctions with Communication," *Journal of Economic Theory*,134, 61-92.

Arrow, K. (1979). "The property rights doctrine and demand revelation under incomplete information," *In Economies and Human Welfare*, Academic Press.

Athey, S., and K. Bagwell (2001). "Optimal Collusion with Private Information," *RAND Journal of Economics*, 32(3), 428-465.

Athey, S., and K. Bagwell (2008). "Collusion with Persistent Cost Shocks," *Econometrica*, 76(3), 493-540.

Athey, S., K. Bagwell, and C. Sanchirico (2004). "Collusion and price rigidity," *Review of Economic Studies*, 71(2), 317-349.

Athey, S., and D. Miller (2007). "Efficiency in repeated trade with hidden valuations," *Theoretical Economics* 2, 299–354.

Athey, S., and I.Segal (2013). "An efficient Dynamic Mechanism,"*Econometrica* 81(6), 2463-2485.

Blume, A., and P. Heidhues, (2002). "Modelling tacit collusion in repeated auctions," mimeo.

d'Aspremont, C., and L.A. Varet-Gerard (1979). "Incentives and Incomplete information," *Journal of Public Economics* 11, 25-45.

Escobar, J. F. and J. Toikka (2013). "Efficiency in Games with Markovian Private Information," *Econometrica* 81, 1887-1934.

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

Hauser, C., and H. Hopenhayn (2008). "Trading Favors: Optimal Exchange and Forgiveness," mimeo.

Hörner, J., and J. Jamison (2007). "Collusion with (Almost) No Information," *RAND Journal of Economics*, 38(3), 804-822.

Jackson, M. O., and H. F. Sonnenschein (2007). "Overcoming Incentive Constraints by Linking Decisions," *Econometrica*, 75(1), 241-258.

Mailath, G., and L. Samuelson (2006). Repeated Games and Reputations. Oxford University Press, Oxford.

Miller D. (2009), "Attainable payofs in repeated games with interdependent private information," mimeo.

Möbius, M. (2001). "Trading Favors," mimeo.

Olszewski W. and M. Safronov (2015). "Efficiency in chip strategies," mimeo.

Radner, R. (1985). "Repeated Principal-Agent Games with Discounting," *Econometrica*, 53(5), 1173-1198.

Spulber, D. F. (1995). "Bertrand Competition When Rivals' Costs Are Unknown," *Journal of Industrial Economics*, 43(1), 1- 11.

Skryzpacz, A., and H. Hopenhayn (2004). "Tacit Collusion in Repeated Auctions," *Journal of Economic Theory*, 114(1), 153-69.