Penalty-card strategies in repeated games

Wojciech Olszewski and Mikhail Safronov*

July 2, 2014

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 penalized by going on suspension; during the period of suspension she obtains a low expected payoff. 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. Players condition their actions only on a simple statistics containing all necessary information regarding the past play. The strategies are also immune to deviations by coalitions, except when the a breakdown of cooperation occurs.

1 Introduction

In the extensive literature on repeated games, games of incomplete information in which types change over time have been relatively less intensively studied and are relatively less understood. This model seems, however, to have a wide range of applications. They include: (a) oligopoly markets in which firms privately know their costs, (b) partnership games in which the effort that can be exerted depends on other duties that partners must perform, or (c) altruistic behavior when the cost of sharing or giving a good depends on a varying utility of consuming it.

In the related 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 (for example, the strategies attain some but not full collusion in oligopoly models), or (b) strategies that allow...
to attain a wider range of outcomes (for example, some kind of folk theorem) but have less intuitive and more complicated form, or have been “tailored” with the objective of attaining particular outcomes.

We study a class of what we call \textit{penalty-card strategy profiles}. The idea is that a player who plays an action resulting in a low expected payoff of other players obtains a penalty card. Of course, it may be in the social interest (e.g., understood as maximizing the aggregate payoff) for a player to play such an action (e.g., this action may yield a very high payoff to the player taking it). So, 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 maximize the aggregate payoff of the other players. Detectable deviations result in a breakdown of cooperation.

Penalty-card strategies resemble what we observe in numerous settings in practice, and have a number of useful features. Players condition their actions only on a simple statistics containing all necessary information regarding the past play; or speaking in more technical terms, the strategy profiles can be implemented by automata with a relatively low number of states, especially in games with a low number of players. The strategies also turn out to be immune to deviations by coalitions, except when a breakdown of cooperation occurs.

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. The fact that the efficient outcome can be attained is intuitive. When the discount factor approaches 1, and players assign higher weights to future payoffs, one can allow a larger limit number of penalty cards. So, players go on suspension less frequently, and the efficient outcome is achieved more frequently.

The property 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 each player internalizes the current payoffs of the opponents by the effect that her actions have on her 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 $\delta$ tending to 1. This allows for omitting all expressions of order $o(\epsilon)$, where $\epsilon = 1 - \delta$, 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, and in this way may

\footnote{Of course, there is here some subtle issue of the kind of order of limits. Inefficient suspensions happen less frequently, but affect payoffs more severely because players become more patient; in addition, they must last longer for providing incentives.}
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. This is, however, not necessary. We prescribe the changes in the structure of penalty cards in the way that each player not only internalizes the effect of her current action on the current payoffs of the opponents, but also internalizes the expected effect of her current action on the future payoffs.

Related literature

Fudenberg, Levine and Maskin (1994) show a folk theorem for a family of repeated games, in which players have i.i.d. types. However, 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 Athey and Bagwell (2001), followed up by Athey, Bagwell and Sanchirico (2004). In their model, firms play a repeated Bertrand game, 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 the model to the more realistic case in which the firms’ costs are more persistent (more precisely, they follows a Markov process in their model). They characterize equilibria in strategies which are independent of the firms’ costs, and construct an equilibrium which depends on the firms’ private information regarding their 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. Escobar and Toikka (2010) show that the efficient payoff vector can be attained in some kind of review strategies; they even prove that any Pareto-efficient payoff vector above a stationary minmax vector can be attained for a more general class of games.

Athey and Bagwell (2001) obtain efficiency in strategies that give the entire market to the firm with the lower cost, and that compensate the firm with the higher cost by promising a larger market share in a future period in which the costs will be equal. These strategies are certainly simple and intuitive, but the authors assume i.i.d. costs and study only duopoly markets. In addition, firms must be able to divide the market in the way they wish if they charge equal prices.

Hörner and Jamison’s (2007) result is general and obtained under weak assumptions, even ones regarding the observability of actions, except the fact that they study only the i.i.d. costs. The strategies are, however, quite involved and carefully “tailored” for obtaining particular payoffs.

Under persistent, Markovian costs, Athey and Bagwell (2008) study strategy profiles such that if a firm declares a lower cost in an odd period, then it takes a larger share of the market in that period, but leaves a larger share of the market for another firm in the next (even) period, independent of their costs. These strategies resemble “one penalty card” strategy profiles. So, not too surprisingly, they achieve only partial, but not full efficiency.
The review strategies from Escobar and Toikka are natural, and deliver general results. One drawback is that they are not entirely explicit. Contingent on some histories, the existence of strategies with some required properties is established by a fixed-point argument, although the chance of reaching such a history is rather low. In addition, our results can be obtained under their assumptions, but also for some class of stage games which violate their assumptions. We will emphasize some other advantages of penalty-card strategies in Section 7.

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. Mobius (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.

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.

2Review 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 is closely related to the linking mechanism from Jackson and Sonnenschein (2007).

3Other related papers include: (a) Skrzypacz and Hopenhayn (2004) who study collusion in auctions and offer a numerical argument that the chip mechanism becomes efficient as discount factor goes to unity; (b) Abdulkadirouglu and Bagwell (2012) who analyze the chip mechanism in a model of favors. (c) Athey and Miller (2007) who look at similar debt strategies in a model of repeated trade with hidden valuations.
2 Preliminaries

2.1 Model

Consider a normal-form game $G$ with $I$ players, numbered by $i = 1, ..., I$. Let $A_i$ and $\Theta_i$ be finite sets of actions and types, respectively, of player $i$. Let $u_i(\theta, a)$ 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 $\delta$. 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 $\Theta_i$. Players also have access to a public randomization device, i.e., they observe the realization of a random variable distributed uniformly on interval $[0, 1]$.

In the main text, we assume that players’ types are i.i.d. according to distributions $\eta_i$, $i = 1, ..., I$. In Appendix, we generalize the results from the main text to players’ types which are still independently distributed, but evolve over time according to time-homogeneous, 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 $\theta'$ 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 $\theta$, the type profile in period $t$ will be $\theta'$ with positive probability. By ergodic theorem, every such process has a limiting type distribution $\eta$, and independent of the initial type profile, 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 repeated-game strategies:

A penalty-card strategy profile has two phases: a cooperation phase, and a joint-penalty phase. At the beginning of each period, if the play is in the cooperation phase, some players may be on suspension. Other players are considered active. Actions prescribed for both active players and players on suspension depend only on types of active players. These actions depend on the history of past play, but only through the set of active players. (Typically, the current play penalizes players on suspension and rewards active players.)

One active player holds a certain number of penalty cards. At the end of each period, this player may obtain another penalty card, or all penalty cards may be annulled, in which case another active player obtains a penalty card. 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.
At the end of each period some players can come back from suspension, and become active. This is independent of the actions of players, just happens when the prescribed suspension comes to its end. It may also happen that at the end of a period the play moves to the joint-penalty phase. Once the joint penalty phase is reached the play remains in this phase forever. The transition rule, that is, the chance of staying in the cooperation phase, the chance of obtaining another penalty card or annulling all existing penalty cards is determined by players’ actions, in a manner that depends on the set of active players.

At the beginning of the repeated game, all players are active, and a randomly drawn player obtains a penalty card.

This class of strategies is sufficient for our purposes. One can, however, consider a more general class of penalty-card strategies. For example, more than one active player may hold penalty cards, only some of the existing penalty cards of a player may be annulled, and not necessarily any player must hold a card. Joint-penalty phase may not be an absorbing state. Or, players’ actions may depend on the current numbers of penalty cards, and the chance of coming back from suspension may depend on the current actions of active players.

The penalty-card strategies can also be viewed as a specific debt contract in which a player holding penalty cards is a borrower, and its debt is jointly owned by active players. Players on suspension are (possibly temporarily) excluded from the credit market, and joint-penalty phase is interpreted as a credit-market failure.

2.3 Assumptions on stage game

We can now present the assumptions that we impose throughout the paper on the stage game $G$. For any set of players $R \subset \{1, \ldots, I\}$ and their type profile $\theta_R \in \Theta_R = \{i \in R : \Theta_i\}$, denote by $a(\theta_R)$ a $R$-efficient action profile, that is, an action profile that maximizes the total payoff of the players from $R$.\footnote{Pick an arbitrary maximizer when there exist more than one.} Let $v_{iR} = E_{\theta_R}(u_i(\theta_i, a(\theta_R)))$ 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 $w_{iR}$ 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 (A): For any $i = 1, \ldots, n$ and $R$ such that $i \notin R$,

$$v_{iR} < w_{R \cup \{i\}}.$$

Assumption (B): For any $i = 1, \ldots, n$ and $R$ such that $i \in R$,

$$\frac{1}{|R| - 1} \sum_{i \neq j \in R} w_{R \setminus \{j\}} > v_{R \setminus \{i\}}.$$
In the cooperation phase of penalty-card strategy profiles, players always play $R$-efficient action profile, for $R$ being the set of currently active players. That is, 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 (C): The incomplete information stage game has an equilibrium in which the payoff of every player $i$ is lower than $w^i_R$ for $R = \{1, ..., I\}$.

We will call the equilibrium described in assumption (C) bad equilibrium. The existence of stage-game equilibria for general games can be established by a simple fixed-point argument. However, some stage-game equilibria may not satisfy assumption (C). There may exist no equilibrium satisfying assumption (C) 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 in 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 (C) is satisfied.

Escobar and Toikka assume the existence of type-independent action profiles, one for each player, that can be used as player-specific punishments. Penalty-card strategies can be easily adapted to approximate efficient payoffs under their assumption. The analysis would be even simpler, as we could use the playerspecific punishment when a player reaches the limit number of cards. In addition, with player-specific punishments, assumptions (A)-(C) would be dispensable.

3 The result in the i.i.d. case and the main idea

In the main text, we focus on the i.i.d. types. We will prove that:

**Theorem 1.** If the stage game satisfies assumptions (A)-(C), then the efficient payoff can be approximated in penalty-card equilibria as the discount factor tends to 1.

In Appendix, we generalize Theorem 1 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 in a more specific manner, suppose that players are allowed to make transfers to one another at the end of each period. In addition, we restrict attention here to the case when all players are active.

For any type profile $\theta$, 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 $R = \{1, ..., I\}$.) Consider the following strategies:
(1) Players report their types.

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

(3) For all \( i \neq j \in \{1, \ldots, I\} \), player \( j \) transfers to player \( i \)

\[
E_{\theta_{-i}}(u_j(\theta_j, a(\theta_i, \theta_{-i}))) - E_{\theta_j}E_{\theta_{-j}}(u_j(\theta_j, a(\theta_i, \theta_{-i}))).
\]

This 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 \( \theta'_i \), given truthful reports of other players, is then

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

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

\[
\sum_{j=1}^{I} E_{\theta_{-i}}(u_j(\theta_j, a(\theta'_i, \theta_{-i})))
\]

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.

(4) If any out of equilibrium action is observed, then players switch to playing permanently the bad stage-game Nash equilibrium. This disciplines the players to take action profile \( a(\theta) \) given any report \( \theta \).

4 Efficient repeated-game strategies

In the present section, we specify the details of penalty-card strategies that attain efficiency. First, we introduce some auxiliary terms. Define by

\[
s^j_i = E_{\theta_{-i}}(u_j(\theta_j, a(\theta_i, \theta_{-i}))) - E_{\theta_j}E_{\theta_{-j}}(u_j(\theta_j, a(\theta_i, \theta_{-i})))
\]

the effect of player \( i \)'s report on player \( j \)'s payoff; in particular, \( s^j_i > 0 \) \( (s^j_i \leq 0) \) 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 \( \theta_i \), but we will often disregard its argument as it will cause no confusion. Let

\[
s_i = \sum_{j \neq i} s^j_i
\]

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

Finally, define

\[
p_i = \Pr\{s_i > 0\} \cdot E_{\theta_i}[s_i|s_i > 0],
\]
or equivalently,
\[
p_i = -\Pr\{s_i \leq 0\} \cdot E_{\theta_i}[s_i | s_i \leq 0].
\]

We are now ready to specify the details of our strategies. Denote by F the (active) player who currently holds a positive number of penalty cards. When all penalty cards of player F are annulled, another player obtains a (first) penalty card; denote this player by G. By \(n\) we denote 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. As in the previous section, we first describe the strategies in the case when all players are active.

In period 1, player F is randomly selected, each with probability \(1/I\), and that player begins the game with a penalty card.\(^5\) At the beginning of other periods, F is the player who currently holds a positive number of penalty cards. Player G is selected randomly from the \(I - 1\) players other than F, each of them with probability \(1/(I - 1)\). Suppose that player \(i\) is the current player F and holds \(k\) cards.

1. If F holds fewer than the limit number of penalty cards, and the play is not in joint penalty phase, then players report their types \(\theta\) truthfully and play an efficient action profile \(a(\theta)\).

2. If a 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.

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

3. If \(s_i > 0\), then the penalty-card structure in the following period is determined with probability \(\alpha^i_k\) by F’s (player \(i\)’s) report. If \(s_i \leq 0\), then the penalty-card structure in the following period is determined with probability \(\phi^i_k\) by F’s (player \(i\)’s) report.

Contingent on the penalty-card structure being determined by player \(i\)’s report, if \(s_i > 0\), player \(i\)’s cards are annulled (and player G obtains a card) with probability \(s_i;\)\(^6\) and if \(s_i \leq 0\), player \(i\) obtains another card with probability \(-s_i\). With the remaining probabilities, player \(i\) keeps holding \(k\) cards. Numbers \(\alpha^i_k\) and \(\phi^i_k\) will be specified later.

4. Similarly, if \(s_j > 0\) (\(s_j \leq 0\)), then the penalty-card structure in the following period is determined with probability \(\beta^j_{i,k} (\psi^j_{i,k})\) by G’s (player \(j\)’s) report.

\(^5\)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, which is inconvenient for the analysis of players’ payoffs and incentives, but does not affect the results.

\(^6\)For simplicity, we will assume that \(|s_i| \leq 1\) for every \(i\). If this is not the case, we can normalize terms \(s_i\) by dividing them by \(\max_i |s_i|\), multiplying \(\alpha^i_k\) and \(\phi^i_k\) by the same factor. As it will be clear later, \(\alpha^i_k\) and \(\phi^i_k\) will converge to 0 as the discount factor converges to 1, while terms \(s_i\) do not depend on the discount factor.
Contingent on the penalty-card structure being determined by player $j$’s report, if $s_j \leq 0$, player F’s cards are annulled (and $j$ as player G obtains a card) with probability $-s_j$; and if $s_j > 0$, player F obtains another card with probability $s_j$. With the remaining probabilities, player $i$ keeps holding $k$ cards. Numbers $\beta_{i,k}$ and $\psi_{i,k}$ will be specified later.

(5) With probability

\[
\frac{k}{2(k+1)} - \phi_k p_i - \frac{1}{I - 1} \beta_{i,k} p_j,
\]

player F obtains another penalty card, independent of the actions played (or messages sent) in the current period. With probability

\[
\frac{1}{I - 1} \left( \frac{1}{2(k+1)} - \phi_k p_i - \psi_{i,k} p_j \right),
\]

all current cards are annulled and player G obtains a penalty card, independent of the actions played in the current period. With the remaining probability there is no change in the penalty-card structure, which means that player F keeps holding $k$ penalty cards.\(^7\)

The lottery over which firm is G and the lotteries concerning the penalty-card structure in the following period are performed at the end of the period (or equivalently, at the beginning of the next period) by using the public randomizing device.

**Remark 1** Note that under these strategies the expected change of in the number of cards of player F is 0: player F obtains one more card with probability $k/2(k+1)$ and all her $k$ 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 $v^i = v^i_R$ for $R = \{1, \ldots, I\} \setminus \{i\}$ be the stage game payoff of player $i$ when she is the only player on suspension, and let $w^i_j = w^i_{R}$ for $R = \{1, \ldots, I\} \setminus \{j\}$ be the stage game payoff of player $i$ when some other player $j$ is the only player on suspension, and let $w^i = w^i_{\{1, \ldots, I\}}$ be the stage game payoff of player $i$ when all players are active. Denote by $V^i_k$ the continuation payoff of player $i$ who is currently player F and holds $k$ cards, and by $W^i_{j,k}$ the continuation payoff of player $i$ when $j$ is currently player F and holds $k$ cards. These payoffs are obviously computed assuming that players play the prescribed strategies. We will often call $V^i_k$ and $W^i_{j,k}$ value functions. These functions are payoffs at an ex ante stage, when players have not yet learned their current types.

\(^7\)Numbers $\alpha_k, \phi_k, \beta_k, \psi_k$ will be small, so the formulas in the displays define positive numbers.
Numbers $\alpha^i_k$ and $\phi^i_k$, which will sometimes be called *probabilities of control*, will be defined so that the following equalities are satisfied:

$$\alpha^i_k s_i (1 - \varepsilon) \left[ \frac{1}{I - 1} W^i_{j,k} - V^i_k \right] = s_i \varepsilon$$

for $s_i > 0$, and

$$\phi^i_k (-s_i) (1 - \varepsilon) [V^i_{k+1} - V^i_k] = s_i \varepsilon$$

for $s_i \leq 0$.

This enables us to imitate the strategies described in Section 3 without transfers. Indeed, player $i$’s report affects $s_i$, and in consequence player $i$’s value function in the subgame beginning in the following period. By (3) of the definition of strategies, this effect on the value function is equal to the left-hand sides of these equations. In turn, the right-hand sides are equal to the sum of stage-game payoffs across all other players, which together with the effect 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.

The two equations are equivalent to

$$\alpha^i_k (1 - \varepsilon) \left[ \frac{1}{I - 1} W^i_{j,k} - V^i_k \right] = \varepsilon \text{and} \phi^i_k (1 - \varepsilon) [V^i_{k+1} - V^i_k] = \varepsilon. \quad (2)$$

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

$$\frac{1}{I - 1} \beta^j_{i,k} s_i (1 - \varepsilon) [W^i_{j,k+1} - W^i_{j,k}] = s_i \varepsilon$$

for $s_i > 0$, and

$$\frac{1}{I - 1} \psi^j_{i,k} (-s_i) (1 - \varepsilon) [V^i_{j,k} - W^i_{j,k}] = s_i \varepsilon$$

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 are equivalent to

$$\frac{1}{I - 1} \beta^j_{i,k} (1 - \varepsilon) [W^i_{j,k+1} - W^i_{j,k}] = \varepsilon \text{and} \frac{1}{I - 1} \psi^j_{i,k} (1 - \varepsilon) [W^i_{j,k} - V^1_i] = \varepsilon. \quad (3)$$

Given the prescribed strategies, value $V^i_k$ for $k = 1, \ldots, n - 1$ satisfies the following recursive equation:

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

Indeed, player $i$’s current stage game payoff is $w^i$. By (3) 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 (4) 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. Together with the chance of obtaining another card described in (5) 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 of the number of cards of player \( i \) being unaltered.

Moving the term \((1 - \varepsilon)V_k^i/2\) to the left-hand side, then dividing the equation by \(1 - (1 - \varepsilon)/2\), and omitting all terms of order smaller than \(\varepsilon\), one can rewrite the recursive equation as:

\[
V_k^i = 2\varepsilon w^i + (1 - 2\varepsilon)\frac{k}{k + 1} V_{k+1}^i + (1 - 2\varepsilon) \frac{1}{I - \frac{1}{1 - \varepsilon}} \frac{1}{n} \frac{1}{j \neq i} W_{j,1}^i,
\]

and for \( k = n \) (also omitting terms smaller than \(\varepsilon\)):

\[
V_n^i = M\varepsilon v^i + (1 - M\varepsilon) V_{n-1}^i.
\]

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, we will see that this may not be entirely true. The equation will be satisfied only in approximation. We will return to this issue in Section 6.

Similarly, value \( W_{j,k}^i \) for \( k = 1, \ldots, n - 1 \) satisfies the following recursive equation:

\[
W_{j,k}^i = 2\varepsilon w^i + (1 - 2\varepsilon)\frac{k}{k + 1} W_{j,k+1}^i + (1 - 2\varepsilon) \frac{1}{I - \frac{1}{1 - \varepsilon}} \frac{1}{n} \frac{1}{m \neq i, j} W_{m,1}^i + (1 - 2\varepsilon) \frac{1}{I - \frac{1}{1 - \varepsilon}} \frac{1}{n} \frac{1}{j \neq i} W_{i,j}^1,
\]

and for \( k = n \):

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

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 (??) and (??).

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 \(\varepsilon\) is very small, so one can disregard terms of order lower than \(\varepsilon\).

From the two formulas for \( V_k^i \), we can derive the expression for \( V_n^i \) as a function of \( W_{j,1}^i \) as follows:

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

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

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

\(^8\)When we omit terms of order lower than \(\varepsilon\), then dividing by \(1 - (1 - \varepsilon)/2 = 1/2 + \varepsilon/2\) is equivalent to multiplying by \(2 - 2\varepsilon\).
which yields
\[ V_{n-1}^i = 2n \varepsilon w^i + (n - 1) M \varepsilon v^i + \{1 - [2n + (n - 1) M] \varepsilon \} \frac{1}{I - 1} W_{j,1}^i. \]

We will now recursively demonstrate that
\[ V_k^i = \varepsilon C_k^i + (1 - D_k \varepsilon) \frac{1}{I - 1} W_{j,1}^i \]
for some constants \( C_k^i \) and \( D_k \).

Suppose (5) holds for \( k + 1 \). Since
\[ V_k^i = 2 \varepsilon w^i + (1 - 2 \varepsilon) \frac{k}{k + 1} \left[ \varepsilon C_{k+1}^i + (1 - D_{k+1} \varepsilon) \frac{1}{I - 1} W_{j,1}^i \right] + (1 - 2 \varepsilon) \frac{1}{I - 1} \frac{1}{k + 1} W_{j,1}^i \]
\[ = \varepsilon \left( 2 w^i + \frac{k}{k + 1} C_{k+1}^i \right) + \left[ 1 - \varepsilon \left( 2 + \frac{k}{k + 1} D_{k+1} \right) \right] \frac{1}{I - 1} W_{j,1}^i, \]
(5) holds for \( k \), and
\[ C_k^i = 2 w^i + \frac{k}{k + 1} C_{k+1}^i \text{ and } D_k = 2 + \frac{k}{k + 1} D_{k+1}. \]

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
\[ U_{j,1}^i = \frac{1}{I - 1} W_{m,1}^i + \frac{1}{I - 1} V_1^i; \]
and
\[ W_{j,k}^i = \varepsilon c_{j,k}^i + (1 - d_k \varepsilon) U_{j,1}^i, \]
where
\[ c_{j,k}^i = 2 w^i + \frac{k}{k + 1} c_{j,k+1}^i \text{ and } d_k = 2 + \frac{k}{k + 1} d_{k+1}. \]

By summing up (5) for \( k = 1 \) across all \( j \neq i \), we obtain
\[ \frac{1}{I - 1} W_{j,1}^i = \varepsilon j_{j,i}^i c_{j,1}^i + [1 - (I - 1) d_1 \varepsilon] V_1^i. \]
This equation together with (5) for \( k = 1 \) enables us to compute \( V_1^i \):
\[ V_1^i = \frac{C_1^i + j_{j,i} c_{j,1}^i}{D_1 + (I - 1) d_1}, \]
and
\[ \frac{1}{I - 1} W_{j,1}^i = \frac{1 + (D_1 \varepsilon) j_{j,i} c_{j,1}^i + [1 - (I - 1) d_1 \varepsilon] C_1^i}{D_1 + (I - 1) d_1}. \]

By recursive computing, we obtain:
\[ C_1^i = 2 w^i \left( 1 + \frac{1}{2} + \ldots + \frac{1}{n - 2} \right) + \frac{2 n w^i}{n - 1} + M v^i, \]
\[ D_1 = 2 \left( 1 + \frac{1}{2} + \ldots + \frac{1}{n - 2} \right) + \frac{2 n}{n - 1} + M, \]
which yields
\[ V_{n-1}^i = 2n \varepsilon w^i + (n - 1) M \varepsilon v^i + \{1 - [2n + (n - 1) M] \varepsilon \} \frac{1}{I - 1} W_{j,1}^i. \]
\[ c_{j,1}^i = 2w^i \left( 1 + \frac{1}{2} + \ldots + \frac{1}{n-2} \right) + \frac{2nw^i}{n-1} + Mw_j^i, \]

and

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

This yields

\[ \lim_{n \to \infty} V_{1}^i = w^i \quad (8) \]

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

\[ \prod_{m=1}^{n} \frac{1}{m}. \]

Similarly, we obtain that

\[ \lim_{n \to \infty} \frac{1}{I - 1_{j \neq i}} W_{j,1}^i = w^i. \quad (9) \]

### 5.3 Probabilities of control, incentives

The probabilities of control \( \alpha_k^i, \phi_k^i, \beta_k^i \) and \( \psi_k^i \) are defined by conditions (??) and (??), where the value functions are determined by the recursive equations from the previous section. It only remains to show that these numbers are positive but small. We will show that this is the case for \( \alpha_k^i \) and \( \phi_k^i \); for \( \beta_k^i \) and \( \psi_k^i \), the argument is analogous.

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

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

and

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

By (??)

\[ V_k^i - \frac{1}{I - 1_{j \neq i}} W_{j,1}^i = \varepsilon C_k^i - D_k \varepsilon - \frac{1}{I - 1_{j \neq i}} W_{j,1}^i, \]

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

\[ V_1^i = w^i + \frac{M}{2 \left( 1 + \frac{1}{2} + \ldots + \frac{1}{n-2} \right) + \frac{2n}{n-1} + M} \left( \frac{1}{I} v_i + \frac{1}{I_{j \neq i}} w_j^i - w^i \right). \quad (10) \]

If the expression in parenthesis is negative, then

\[ V_k^i - \frac{1}{I - 1_{j \neq i}} W_{j,1}^i = \varepsilon M (v^i - w^i) - \varepsilon \frac{2 \left( 1 + \frac{k}{k+1} + \frac{k}{k+2} + \ldots + \frac{k}{n-2} \right) + \frac{2kn}{n-1} + kM}{2 \left( 1 + \frac{1}{2} + \ldots + \frac{1}{n-2} \right) + \frac{2n}{n-1} + M} \left( \frac{1}{I} v_i + \frac{1}{I_{j \neq i}} w_j^i - w^i \right) \]
\[ \leq \epsilon k M (v^i - w^i) - \epsilon k M \left( \frac{1}{I} v^i + \frac{1}{I \neq i} \sum_{j \neq i} w^i_j - w^i \right) = -\epsilon k M \frac{I-1}{I} \left( \frac{1}{I-1 \neq i} \sum_{j \neq i} w^i_j - v^i \right), \]

and by assumption (B), this is a negative expression at least of order \( kM \epsilon \), which implies that \( \alpha^i_k \) is positive but small if \( M \) is sufficiently large.

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

\[ V^i_k - \frac{1}{I-1 \neq i} \sum_{j \neq i} W^i_{j,1} < \epsilon k M (v^i - w^i), \]

and so by assumption (A), this is again a negative expression at least of order \( kM \epsilon \).

Similarly,

\[ V^i_k - V^i_{k+1} = \epsilon [C^i_k - C^i_{k+1}] - [D_k - D_{k+1}] \epsilon \frac{1}{I-1 \neq i} \sum_{j \neq i} W^i_{j,1}, \]

which by the recursive equations for \( C^i_k \) and \( D_k \), and (??) is equal to

\[ -\epsilon \frac{1}{k+1} \left( C^i_{k+1} - D_{k+1} \frac{1}{I-1 \neq i} \sum_{j \neq i} W^i_{j,1} \right) + 2 \epsilon \left( v^i - \frac{1}{I-1 \neq i} \sum_{j \neq i} W^i_{j,1} \right) \]

\[ = -\frac{1}{k+1} \left( V^i_{k+1} - \frac{1}{I-1 \neq i} \sum_{j \neq i} W^i_{j,1} \right) + 2 \epsilon \left( v^i - \frac{1}{I-1 \neq i} \sum_{j \neq i} W^i_{j,1} \right). \]

So, by the previous argument and (??), we also have that \( \phi^i_k \) 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, a player on suspension is allowed to come back and become active every period with probability \( \mu \) such that

\[ M = \sum_{t=1}^{\infty} t \mu (1 - \mu)^{t-1} = \frac{1}{\mu}. \]

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 game, 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). If player \( i_1 \) goes on suspension again, then players start playing the \( (I-1) \)-player game from the beginning, not from the moment they stopped because player \( i_1 \) returned from suspension.

It may happen in the \( (I-1) \)-player game 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 from the $I$-player game, and when she returns, all players who went on suspension later return as well.

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 game. If player $i_k$ returns from suspension, then all players $i_{k+1}, ..., i_l$ return as well. Moreover, and if it ever happens again that players $i_1, i_2, ..., i_l$ are on suspension, the $(I - l)$-player subgame is played from the beginning, without returning the past play in this subgame.

From the perspective of the $(I - l)$-player subgame, the probability of interrupting the subgame, when one of the players $i_1, i_2, ..., i_{l-1}$ 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 whose existence was assumed in (C), even if this phase begins when some players are on 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 $V^i_n$ 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 $w^j_i$. This, however, is 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 $\delta$ 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 they can be implemented by automata with a relatively low number of states, especially in games with a low number of players. In other words, players condition their actions only on a simple statistics containing all necessary information regarding the past play. The states can be described by whether the play is in the cooperation or joint-penalty phase, who are on suspension and the order in which players went on suspension, and who currently holds penalty cards and how many of them.

In particular, the states of the penalty-card strategy profiles that approximate efficiency only minimally
Therefore, penalty-card strategies seem particularly attractive in games with a small number of players, and large numbers of types and actions. In contrast, if a review strategy profile prescribes different actions for different types, and the number of both types and actions is large, then playing 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 Coalition proofness

The strategies used to prove Theorem 1 have another important feature, which is not a feature of all penalty-card strategy profiles. Namely, those strategies are coalition-proof on equilibrium paths. More precisely, it can be readily checked that those strategies have the following property:\footnote{This is simply a property of the AGV mechanism.} Suppose that for some reason once the joint-penalty phase is reached, no coalition of players is allowed to deviate. Then, if the play is in cooperation phase, no coalition of players is able to Pareto-improve their payoffs by deviating jointly. In other words, if the play is in cooperation phase, no coalition of players can Pareto-improve their payoffs by deviating jointly in the current period and returning to the prescribed strategies in the subsequent periods.

It is unclear whether a similar property could be obtained by means of review strategies. The problem is somewhat similar to that pointed out by Escobar and Toikka when types are Markov instead of i.i.d. Namely, Escobar and Toikka pointed out that if players are tested only for the right frequency of their actions, players with Markovian types may benefit from correlating their reports over time; similarly, some players can benefit from playing as a coalition and reporting their types jointly. A coalition of players may correlate their reports, maintaining the property that each player’s reports will approximate that player’s probability distribution over types.

Of course, one can try to modify of the linking mechanism of Jackson and Sonnenschein, or the mechanism in Escobar and Toikka, by including to the review blocks tests against deviations by coalitions. However, to verify that this delivers coalition-proof equilibria on equilibrium paths with efficient payoffs would require the amount of work similar to that performed in Escobar and Toikka in order to adjust the linking mechanism to the situations in which types are Markovian. In addition, this would require conducting in all review blocks a number of tests that is exponential in the number of players, one for every subset of players.

\footnote{The number of states of the “efficient” penalty-card strategy profiles may depend on the space of types and actions in the following way: As we enlarge the space, we may create new possibilities for profitable deviations from efficient actions. This is manifested in assumptions (A) and (B) by smaller margins by which the two inequalities hold. This, in turn, requires increasing the limit on the number of penalty cards that players are allowed to hold, for our strategies to be correctly defined. (See, for example, the argument showing that the probabilities of control are positive but small.)}
7.3 Other advantages, and limitations

The analysis of the previous section, and that of Escobar and Toikka suggests that penalty-card strategies, or more generally “debt strategies” dominate the review strategies, in the sense that any test can induce only some specific features of behavior, in specific settings, while debt is a more universal way of providing incentives. More specifically, tests that induce desired behavior when types are i.i.d. may fail when their types are Markov, or we need to include specific tests when we wish to prevent coalitional deviations. In turn, penalty card (or debt) strategies allow players to use, within some limits, their private information in the way that is most beneficial for them.

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 were able to show in a working paper version of the present paper that the same penalty-card strategies approximate efficiency in the repeated Spulber’s oligopoly game, studies by Athey and Bagwell and others, for all Markov chains with transition probability bounded away from 1.

In terms of limitations, the analysis 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, and coalitional deviation is not an essential issue. In addition, at least the penalty-card strategies used to prove our general results rely heavily, and in a subtle manner, on public randomization. However, public randomization is dispensable in some more specific settings; for example, we show in a working paper version of the present paper that it is dispensable in the repeated Spulber’s duopoly game.

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 in a desirable or not desirable way, giving players 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 report and their ex ante 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

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

(12)
where \( \theta^{-1} = (\theta^{-1}_i, \theta^{-1}_-i) \) denotes the type profile reported in the previous period, and \( E_{\theta_{-i}}[\cdot | \theta^{-1}_-i] \) and \( E_\theta[\cdot | \theta^{-1}] \) mean that the expected values over \( \theta_{-i} \) and \( \theta \) are computed given the distribution of \( \theta_{-i} \) and \( \theta \), respectively, determined by the previous reports \( \theta^{-1}_-i \) and \( \theta^{-1} \).

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 \( \theta^{-1}_i \) on her continuation payoff (beginning in the current period) would be \( \varepsilon \) times expression (??). The first term of this expression does not depend on report \( \theta^{-1}_i \), it does depend only on report \( \theta_i \). The second term, in turn, is determined before observing the current report \( \theta_i \), and depends on player \( i \)'s previous report \( \theta^{-1}_i \).

This means that player \( i \)'s report \( \theta^{-1}_i \) affects the value of \( s_i \) not only in the period it was sent, but also in the following period. In other words, player \( i \) has an additional incentive, compared to the i.i.d. case, to report \( \theta^{-1}_i \)'s that give low values of \( E_\theta[u_j(\theta_j, a(\theta_i, \theta_{-i})) | \theta^{-1}] \). 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 \( s_i \):

\[
(1 - \varepsilon)_{j \neq i} E_{\theta^{i+1}}[u_j(\theta^{i+1}_j, a(\theta^{i+1}_i, \theta^{i-1}_-i)) | \theta_i, \theta^{-1}_-i],
\]

where \( \theta^{+1} = (\theta^{+1}_i, \theta^{+1}_-i) \) denotes the next-period type profile. This term removes the additional incentives. However, a new problem appears, namely, the expected value of all these new terms, given \( \theta^{-1} \), may not be equal to 0. And this would make the probability distribution over penalty cards in the following period depend on the currently reported type profile. This 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. This, in turn, creates again additional incentives. As a result, one keeps including newer terms to the formula for \( s_i \). These newer terms refer to the expectation of what will happen in more remote future given the current report. Due to the convergence of Markov chain to the limiting distribution 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 \( s_i \) an expression of order \( O(\varepsilon) \), while an infinite number of terms would make \( s_i \) 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.

### 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 \( \eta \). Define \( \pi_R \) and \( \pi_R^i \) 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 $\eta$. We make analogous assumptions, to the i.i.d. case:

Assumption (A'): For any $i = 1, \ldots, n$ and $R$ such that $i \notin R$,

$$\tau_{R}^i < \tau_{R \cup \{i\}}^i.$$  

Assumption (B'): For any $i = 1, \ldots, n$ and $R$ such that $i \in R$, we have that

$$\frac{1}{|R| - 1} \sum_{j \in R \setminus \{i\}} \tau_{R \setminus \{j\}}^i > \tau_{R \setminus \{i\}}^i.$$  

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

We elaborate on assumption (C') 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 (D): For all type profiles $\theta \in \Theta$, and all subsets $R \subset \{1, \ldots, I\}$ 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 1 for Markov types:

**Theorem 2.** If assumptions (A')-(C') and (D) are satisfied, then the efficient payoff can be approximated in penalty-card equilibria as the discount factor tends to 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 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 (C') is satisfied. Therefore, it must be assumed that there exist equilibria in which every player obtains a lower payoff than in the efficient outcome.
This might suggest that assumption (C’) is difficult to check, making our result in the Markovian case not too useful. This is not true. It is relatively easy to construct explicitly some “bad” repeated-game equilibria in many concrete settings (such as the repeated version Spulber’s of oligopoly game). In addition, if one is interested in symmetric games, then Theorem 2 delivers efficient strategies by an argument analogous to that from Section 2.3.

9 Appendix

The purpose of this appendix is to prove Theorem 2.

9.1 Efficient strategies

The strategies will be similar to i.i.d. case. 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 penalty cards. Player G is selected randomly from the $I-1$ players other than F, each of them with probability $1/(I-1)$. Suppose that player F currently holds $k$ cards.

(1) If F holds fewer than the limit number of penalty cards $n$, and the play is not in joint penalty phase, then players report their types $\theta = \theta_R$ for $R = \{1, \ldots, I\}$ truthfully, and play the efficient action profile $a(\theta)$.

(2) 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 $R = \{1, \ldots, I\} - \{i\}$. When player $i$ comes back from suspension, one penalty card of player $i$ is annulled.

This prescription of play applies only under the assumption that no other player goes on suspension during the suspension of player $i$. We will specify later the details of play when a player goes on suspension during the suspension of player $i$. It will be important that $M$ and $n$ diverge to infinity at the rates such that

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

(3) With probability $\alpha^i_k$ or $\phi^i_k$ the penalty-card structure in the following period is determined by F’s (player i’s) report, depending on whether $s^i_k(\theta^{-1}, \theta_i) > 0$ or $s^i_k(\theta^{-1}, \theta_i) \leq 0$, respectively. (We will define $\alpha^i_k$, $\phi^i_k$ and $s^i_k(\theta^{-1}, \theta_i)$ later.)

Contingent on the penalty-card structure being determined by player i’s report, if $s^i_k(\theta^{-1}, \theta_i) > 0$, player i’s cards are annulled (and player G obtains a card) with probability $s^i_k(\theta^{-1}, \theta_i)$; and if $s^i_k(\theta^{-1}, \theta_i) \leq 0$, player i obtains another card with probability $-s^i_k(\theta^{-1}, \theta_i)$. With the remaining probability, player i keeps holding $k$ cards.

(4) With probability $\beta^j_{i,k}$ and $\psi^j_{i,k}$, the penalty-card structure in the following period is determined by G’s (player j’s) report, depending on whether $s^j_{i,k}(\theta^{-1}, \theta_j) > 0$ or $s^j_{i,k}(\theta^{-1}, \theta_j) \leq 0$, respectively. (We will
define $\beta^j_{i,k}, \psi^j_{i,k}$ and $s^j_{i,k}(\theta^{-1}, \theta_j)$ later.)

Contingent on the penalty-card structure being determined by player $j$’s report, if $s^j_{i,k}(\theta^{-1}, \theta_j) \leq 0$, player $F$’s cards are annulled (and $j$ as player $G$ obtains a card) with probability $-s^j_{i,k}(\theta^{-1}, \theta_j)$; and if $s^j_{i,k}(\theta^{-1}, \theta_j) > 0$, player $F$ obtains another card with probability $s^j_{i,k}(\theta^{-1}, \theta_j)$. With the remaining probability, player $i$ keeps holding $k$ cards.

(5) With probability
\[ \frac{k}{2(k+1)} - \phi^j_k p^j_k(\theta^{-1}) - \beta^j_{i,k} \frac{1}{T-1} \sum_{j \neq i} p^j_{i,k}(\theta^{-1}), \]
player $F$ obtains another penalty card, independent of the actions played (or messages sent) in the current period, and with probability
\[ \frac{1}{T-1} \left( \frac{1}{2(k+1)} - \alpha^j_k p^j_k(\theta^{-1}) - \psi^j_{i,k} p^j_{i,k}(\theta^{-1}) \right), \]
all current cards are annulled and player $G$ obtains a penalty card. (Again, $p^j_k(\theta^{-1})$ and $p^j_{i,k}(\theta^{-1})$ will be defined later.)

With the remaining probabilities, $F$ keeps holding $k$ penalty cards.

As in the i.i.d. case, the lottery over which player is $G$ and the lotteries determining the penalty-card structure in the following period are performed at the end of the period by using the public randomization device. Also, if a player does not play in the cooperation phase the prescribed action $a(\theta)$ (or $a_R(\theta)$ for a proper subset of players $R$ when players from $\{0, ..., I\} - R$ are on suspension), the play switches to the joint-penalty phase in which players play the equilibrium described in assumption $(C')$.

### 9.2 Missing definitions

We define the probabilities of control $\alpha^i_k, \phi^i_k, \beta^i_{j,k}$, and $\psi^i_{j,k}$ by equations (??) and (??), where value function $V^i_k$ and $W^i_{j,k}$ are as in the i.i.d. case with the limit $\eta$ 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:

\[ \alpha^i_k = \frac{A^i_k}{M(k+1)} , \beta^i_{j,k} = \frac{B^i_{j,k}}{M^2} , \phi^i_k = \frac{\Phi^i_k}{M} , \psi^i_{j,k} = \frac{\Psi^i_{j,k}}{M(k+1)}, \]

(13)

where $A^i_k$, $B^i_{j,k}$, $\Phi^i_k$ and $\Psi^i_{j,k} \geq 0$ are bounded by a constant which does not depend on $M$ and $n$.

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

\[ B^i_{k,T}(\theta^{-1}, \theta_i) = \sum_{t=0}^{T} \sum_{j \neq i} (1 - \varepsilon)^t E[u^t_j(\theta_i, \theta^{-1}_i)]. \]

where the expression $E[u^t_j(\theta_i, \theta^{-1}_i)]$ stands for the expectation of the stage-game payoff $u_j$ of player $j$ in $t$ periods from now, given the current type $\theta_i$ of player $i$ and the previous types $\theta^{-1}_i$ of players other than $i$. 

22
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

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

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 $\theta_i$ 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 $\theta_i', \theta_i''$ and $\theta_{-i}$, there exists a number $T$ such that for any $t > T$ we have

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

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 when $t$ grows large. Since players may go on suspension, it may happen that for one type profile, say $\theta_i', \theta_{-i}^{-1}$, the probability that player $i$ will be on suspension $t$ periods from now is high, 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 low. So, the payoffs $u_j^{+t}$ of players $j \neq i$ may be very different. However, the probability of a player 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 $\Delta$ lower than the difference $\sum_{i \in R} [u_i(\theta, a(\theta_R)) - u_i(\theta, a)]$ for all profiles $\theta$, subset of players $R$, and actions $a \neq a(\theta_R)$. By assumption (D) 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 $\theta$. 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 $\theta$, 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 \( \theta \) 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 \( V_{i, \theta}^k \), for \( k < n - 1 \), is equal to the expectation over the type profile \( \theta' \) 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:

\[
(1 - \varepsilon) \frac{1}{I - 1} \sum_{j \neq i} \left[ \alpha_k^i s_k^i(\theta, \theta') \chi \{ s_k^i(\theta, \theta') > 0 \} - \psi_j s_{i,k}^j(\theta, \theta') \chi \{ s_{i,k}^j(\theta, \theta') \leq 0 \} \right] + \left( \frac{1}{2(k + 1)} - \alpha_k^i \right) \frac{1}{I - 1} s_{i,k}^i(\theta) \chi \{ s_{i,k}^i(\theta) \leq 0 \}
\]

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

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

and

\[
(1 - \varepsilon) \left[ \frac{1}{2} + \alpha_k^i \left( p_k^i(\theta) - s_k^i(\theta, \theta') \chi \{ s_k^i(\theta, \theta') \} \right) + \frac{1}{I - 1} \sum_{j \neq i} \psi_j s_{i,k}^j(\theta, \theta') \chi \{ s_{i,k}^j(\theta, \theta') \leq 0 \} \right] + \phi_k(p_k^i(\theta) + s_k^i(\theta, \theta') \chi \{ s_k^i(\theta, \theta') \leq 0 \})
\]

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 happens: (a) when player F’s report determines the penalty-card structure and \( s_k^i(\theta, \theta') > 0 \); in this case, it happens with probability \( \alpha_k^i s_k^i(\theta, \theta') \) (see (3) of the definition of the prescribed strategies); (b) when player G’s report determines the penalty-card structure and \( s_{i,k}^j(\theta, \theta') \leq 0 \); in this case, it happens with probability \( -\psi_j s_{i,k}^j(\theta, \theta') \) (see (4) of the definition of the prescribed strategies); (c) independently of players’ reports with probability given in (5) of the definition of the prescribed strategies.

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

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

\[
V_{i, \theta}^k
\]
where $J^i_{j,k,\theta,\theta'}$, $I^i_{k,\theta,\theta'}$ are terms of order $O(1/M)$ by (??), and because $s^j_k(\theta,\theta'_i)$ is bounded across all values of $\varepsilon$. That is, $V^i_{k,\theta}$ is a sum of a term of order $\varepsilon$, and a weighted average of $W^i_{j,1,\theta'}$, $V^i_{k+1,1,\theta'}$, $V^i_{k,\theta'}$. In addition, we have that

$$\sum_{\theta'} \eta_{\theta,\theta'} J^i_{j,k,\theta,\theta'} = \sum_{\theta'} \eta_{\theta,\theta'} I^i_{k,\theta,\theta'} = 0,$$

so the ex ante probability that $W^i_{j,1,\theta'}$, for some $\theta'$, will be the following period continuation payoff is

$$\frac{1}{1 - 2(k+1)} \frac{1}{I - 1},$$

and the ex ante probability that $V^i_{k+1,1,\theta'}$, for some $\theta'$, will be the following period continuation payoff is $\frac{k}{2(k+1)}$, and with the remaining probability $V^i_{k,\theta'}$, for some $\theta'$, 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 $\eta$,

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

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

and

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

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

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

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

$$+ (1 - \varepsilon) \frac{k}{2(k+1)} I^i_{k,\theta,\theta'} V^i_{k+1,1,\theta'} - (1 - \varepsilon) \frac{1}{2} \left( \frac{1}{k+1} \frac{1}{I - 1} \sum_{j \neq i} J^i_{j,k,\theta,\theta'} + \frac{k}{k+1} I^i_{k,\theta,\theta'} \right) V^i_{k,\theta'}.$$

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 $O(\varepsilon/M)$.

\[11\] By stability we mean that

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

for all $\theta'$. By the ergodic theorem, the limit distribution of any Markov chain has this property.
Claim 2. For any players $i$ and $j$, number of cards $k$, and type profiles $\theta$ and $\theta'$ there is constant $C > 0$, independent of $\varepsilon, M, n$ such that $|V_{k,\theta}^i - V_{k,\theta'}^i| < C\varepsilon$ and $|W_{j,k,\theta}^i - W_{j,k,\theta'}^i| < C\varepsilon$. Analogous estimates hold for $\hat{V}_{n-1,\theta}^i$ and $\hat{W}_{j,n-1,\theta'}^i$.

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

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

If we disregard the terms of order lower than $\varepsilon$, we can transform this formula into

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

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

It appears that this formula for $\hat{V}_k^i$, differs from the formula for $V_k^i$ in the i.i.d. case only by replacing $\bar{w}^i$ with $\bar{w}_k^i$. One can perform similar calculations for $\hat{W}_{j,k}^i$, $\hat{V}_{n-1,k}^i$, and $\hat{W}_{j,n-1,k}^i$, and then repeat the reasoning from the i.i.d. case to obtain that

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

This, together with Claim 2 implies that our penalty-card strategies attain, as the discount factor tends to 1, the efficient payoffs.

9.4 Incentives

Observe first that $\hat{V}_k^i$ and $\hat{W}_{j,k}^i$ 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 $\bar{w}^i$ replaced with $\bar{w}_k^i$, and the differences between $\bar{w}^i$ and $\bar{w}_k^i$ are of order $O(1/M)$. Therefore, by the same calculations as in the i.i.d. case (see formulas (??) and (??)), we obtain that

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

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

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

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

12This 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 $\theta$ and $\theta'$. 

26
Denote the current actual type profile as \( \theta^* \). By inspection of the formula for \( V_{k,t}^i \), one can see that player \( i \)'s current report \( \hat{\theta}_i^t \) affects: the current payoff \( w_i^{(\theta^*, \hat{\theta}_i^t)} \), the value of \( s_k^i(\theta, \hat{\theta}_i^t) \), and continuation payoffs \( W_{j,1,(\theta^*, \hat{\theta}_i^t)}^i \) and \( V_{k,(\theta^*, \hat{\theta}_i^t)}^i \). More specifically, \( s_k^i(\theta, \hat{\theta}_i^t) \) is affected through its first component \( B_{k,T}^i(\theta_{-i}, \hat{\theta}_i^t) \), and the continuation payoffs are affected through the value of \( E_{\theta_i^{t+1}}[B_{k,T}^i(\theta_{-i}, \theta_i^{t+1}) | \hat{\theta}_i^t] \).

We will first estimate the effect of player \( i \)'s report \( \hat{\theta}_i^t \) on \( W_{j,1,(\theta^*, \hat{\theta}_i^t)}^i \). 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 \( \theta_i^t \), rather than the reported type \( \hat{\theta}_i^t \).

If the future report \( \theta_i^{t+1} \) is such that \( B_{j,1,T}^i(\theta_{-i}^t, \theta_i^{t+1}) > E_{\theta_i^{t+1}}[B_{j,1,T}^i(\theta_{-i}^t, \theta_i^{t+1}) | \hat{\theta}_i^t] \), then \( W_{j,1,(\theta^*, \hat{\theta}_i^t)}^i \) depends on player \( i \)'s report \( \hat{\theta}_i^t \) through

\[
(B_{j,1,T}^i(\theta_{-i}^t, \theta_i^{t+1}) - E_{\theta_i^{t+1}}[B_{j,1,T}^i(\theta_{-i}^t, \theta_i^{t+1}) | \hat{\theta}_i^t]) \frac{1}{I-1} \beta_{1,1}^{i,j}[W_{j,1,0}^{i} - W_{j,1,\theta_i^{t+1}}^i]
\]

The first equality follows from the fact that \( W_{j,1,\theta_i^{t+1}} = \hat{W}_{j,1}^i + O(\varepsilon) \), 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 \( \beta_{1,1}^i = O(1/M) \). Similarly, if \( B_{j,1,T}^i(\theta_{-i}^t, \theta_i^{t+1}) \leq E_{\theta_i^{t+1}}[B_{j,1,T}^i(\theta_{-i}^t, \theta_i^{t+1}) | \hat{\theta}_i^t] \), then \( W_{j,1,(\theta^*, \hat{\theta}_i^t)}^i \) depends on player \( i \)'s report \( \hat{\theta}_i^t \) through

\[
(B_{j,1,T}^i(\theta_{-i}^t, \theta_i^{t+1}) - E_{\theta_i^{t+1}}[B_{j,1,T}^i(\theta_{-i}^t, \theta_i^{t+1}) | \hat{\theta}_i^t]) \frac{1}{I-1} \psi_{1,1}^{i,j}[V_{j,1,0}^{i} - \hat{V}_{j,1}^i]
\]

The effects of player \( i \)'s report \( \hat{\theta}_i^t \) on \( V_{k+1,(\theta^*, \hat{\theta}_i^t)}^i \) and \( V_{k,(\theta^*, \hat{\theta}_i^t)}^i \), take the same form. The overall effect of report \( \hat{\theta}_i^t \) on player \( i \)'s value function through continuation payoffs must be adjusted by factor \((1 - \varepsilon)\), and considered in expectation contingent on \( \theta_{-i}^t \). This yields

\[-(1 - \varepsilon)E_{\theta_{-i}^t}E_{\theta_i^{t+1}}[B_{j,1,T}^i(\theta_{-i}^t, \theta_i^{t+1}) | \hat{\theta}_i^t, \theta_{-i}^t](\varepsilon + O(\varepsilon/M)).\]

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

\[
B_{k,T}^i(\theta_{-i}, \hat{\theta}_i^t)(\varepsilon + O(\varepsilon/M)) - (1 - \varepsilon)E_{\theta_{-i}^t}E_{\theta_i^{t+1}}[B_{j,1,T}^i(\theta_{-i}^t, \theta_i^{t+1}) | \hat{\theta}_i^t, \theta_{-i}^t](\varepsilon + O(\varepsilon/M))
\]

\[
= \varepsilon[B_{k,T}^i(\theta_{-i}, \hat{\theta}_i^t) - (1 - \varepsilon)E_{\theta_{-i}^t}E_{\theta_i^{t+1}}[B_{j,1,T}^i(\theta_{-i}^t, \theta_i^{t+1}) | \hat{\theta}_i^t, \theta_{-i}^t]] + O(\varepsilon/M).
\]

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

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

where first term is a sum over all players but \(i\), as none is on suspension, while second term includes the possibility that some player is on suspension. The first of the two sums is equal to

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

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 \(\Delta\), 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:

1. When some players are on suspension, values of \(B_{i,k,T}^n(\theta_{i-1}^{-1}, \theta_i)\) and \(s_{i,k}^n(\theta_{i-1}^{-1}, \theta_i)\) 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.

2. Recall that \(s_{i,n-1}^n(\theta_{i-1}^{-1}, \theta_i) = B_{n-1,T}^i(\theta_{i-1}^{-1}, \theta_i) - E_{\theta_i}[B_{n-1,T}^i(\theta_{i-1}^{-1}, \theta_i) | \theta_{i-1}^{-1}]\). It is important that when \(s_{i,n-1}^n(\theta_{i-1}^{-1}, \theta_i)\) is computed for the period following a return from suspension, \(\theta_{i-1}^{-1}\) in its second term \(E_{\theta_i}[B_{k,T}^i(\theta_{i-1}^{-1}, \theta_i) | \theta_{i-1}^{-1}]\) 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 \(E_{\theta_i}[B_{k,T}^i(\theta_{i-1}^{-1}, \theta_i) | \theta_{i-1}^{-1}]\) in the case she becomes active next period. A similar comment applies to \(s_{i,k}^2(\theta_{i-1}, \theta_{i+1})\).

3. We estimated the impact of player \(i\)’s report \(\hat{\theta}_i\) on \(W_{j,1,1}(\theta_{j+1}, \theta_j)\) \(V_{j+1,1}(\theta_{j+1}, \theta_j)\) and \(V_{j,2}(\theta_{j+1}, \theta_j)\) 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 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


