

When the Punishment Must Fit the Crime: Remarks on the Failure of Simple Penal Codes in Extensive-Form Games*

George J. Mailath[†] Volker Nocke[‡]
University of Pennsylvania University of Pennsylvania and CEPR
Lucy White[§]
Harvard Business School and CEPR

First version August 2004
March 28, 2006

Abstract

In repeated normal-form games, *simple penal codes* (Abreu 1986, 1988) permit an elegant characterization of the set of subgame-perfect outcomes. We show that the logic of simple penal codes fails in repeated *extensive*-form games. We provide two examples illustrating that a subgame-perfect outcome may be supported *only* by a profile with the property that the continuation play after a deviation is tailored not only to the identity of the deviator, but also to the nature of the deviation.

Keywords: Simple Penal Code, Subgame Perfect Equilibrium, Repeated Extensive Game, Optimal Punishment

JEL Classification Codes: C70, C72, C73.

*Mailath and Nocke thank the National Science Foundation for research support (grants SES-0350969 and SES-0422778). We thank Larry Samuelson for helpful discussions.

[†]Department of Economics, University of Pennsylvania, 3718 Locust Walk, Philadelphia PA 19104, USA. Email: gmailto@econ.upenn.edu. URL: www.econ.upenn.edu/~gmailto/.

[‡]Department of Economics, University of Pennsylvania, 3718 Locust Walk, Philadelphia PA 19104, USA. Email: nocke@econ.upenn.edu. URL: www.econ.upenn.edu/~nocke/.

[§]Harvard Business School, Soldiers Field Road, Boston MA 02163, USA. Tel: (+1) 617 495 0645. Fax: (+1) 617 496 7357. E-mail: lwhite@hbs.edu.

1. Introduction

Characterizing the set of equilibrium payoffs in repeated games is often a subtle and difficult task. The folk theorem for perfect-monitoring repeated games (Fudenberg and Maskin, 1986) assures us only that every feasible and strictly individually-rational payoff is, under relatively weak conditions, the payoff of some subgame-perfect equilibrium, *provided* players are sufficiently patient. On the other hand, we are often interested in situations in which players are not arbitrarily patient, where the folk theorem is less helpful. For repeated *normal*-form games, the *simple penal codes* of Abreu (1986, 1988) provide an elegant approach to the problem: Discounting guarantees that any subgame-perfect outcome can be supported by a profile with the property that any deviation by a player from the current prescribed path is punished by the same punishment path (penal code), i.e., the continuation play after a deviation by a player is *independent* of the nature of the deviation, depending only on the identity of the deviator.

In this paper, we consider repeated *extensive*-form games. In each *period*, players play the extensive-form game once. Repeated normal-form games (with perfect monitoring) are stationary: after any history, the associated subgame is strategically equivalent to the original repeated game. The situation is more complicated when we consider repeated extensive-form games, since the set of actions available for the player(s) in any given period can depend on the history: different deviations at a given decision node induce different subgames (with potentially different action sets) within the stage game. Consequently, the appropriate notion of simple penal code for a repeated extensive-form game is not obvious. Nonetheless, for any history that ends at the end of a period, the associated subgame is strategically equivalent to the original repeated game. This suggests that any definition of a simple penal code should have the feature that if a player deviates from the candidate equilibrium at a particular node, then subsequent play beginning in the next period should be independent of the precise nature of the deviation at that node.¹ We use two simple examples to show that optimal punishments in repeated extensive-form games need not have this property: even though it would be feasible to play the same outcome path after two different deviations at the same node, sustaining a candidate equilibrium outcome may require that different outcome paths are played in subsequent periods.

Say that an action for a player is myopically suboptimal if, given the specified behavior for the other players, that action is not optimal when payoffs from future periods are ignored. The logic of simple penal codes fails in repeated extensive-form games because some equilibria require the use of within-period myopically-suboptimal “punishments” to ensure that deviations are not profitable. In contrast, in repeated

¹A stricter notion of simplicity requires that whenever a given player deviates (irrespective of the node at which he deviates), the same punishment path must be played. Since simple punishments fail to be optimal in our weaker sense, they also fail to be optimal in this stricter sense.

normal-form games, the sequential rationality of “within-period” punishments is not an issue (see the discussion at the end of Section 3 and in Section 4).

Our examples are once-repeated extensive-form games, since the failure of simple codes already arises in the simpler context of finitely-repeated extensive-form games, and the intuition for the infinitely repeated case is similar. In each example, the within-period punishment is myopically suboptimal for the potential punisher(s), but is required to sustain the desired equilibrium.

In our first example, there are three players, and sustaining the desired equilibrium requires that the two later movers both inflict within-period punishment on the first mover in the event of a deviation. However, there is a conflict of interest between the two potentially enforcing players, so that it is not possible to provide the maximum “reward” to both the punishers in any given equilibrium. Intuitively, therefore, the continuation equilibrium chosen (and hence the “reward” which each punisher receives) must depend on the degree of sacrifice which the punisher makes in inflicting the myopically suboptimal punishment. In other words, the continuation game played after a deviation by the first mover in this example is not simple, but rather depends on *which* deviation he has chosen among those that are available to him.

Our second example focuses on a contrasting force. In our second example, there are only two players so the division of the “carrot” which induces players to apply the “stick” is not an issue. Instead, the difficulty with simple penal codes arises because the interests of the deviator and the potential punisher are aligned: rewarding a player for carrying out the within-period punishment of the earlier deviator necessarily rewards the deviator himself. Consequently, the rewards for the punisher (and hence the outcome path following the deviation) have to be fine-tuned to the particular deviation chosen by the deviator.

Apart from Rubinstein and Wolinsky (1995) and Sorin (1995), the repeated-game literature has focused on repeated normal-form games, ignoring the dynamic structure within the stage game. Rubinstein and Wolinsky (1995) present some examples illustrating the difference between the set of subgame-perfect equilibrium payoffs of repeated extensive and normal-form games for patient players when the standard dimensionality condition does not hold. Sorin (1995) discusses the implications of the different information that players have available across periods in repeated extensive, rather than normal, form games. Neither of these papers is concerned with penal codes.

2. First example: The reward should fit the temptation

“You oughtn’t to yield to temptation.”

“Well somebody must, or the thing becomes absurd.”

Anthony Hope, *The Dolly Dialogues*.

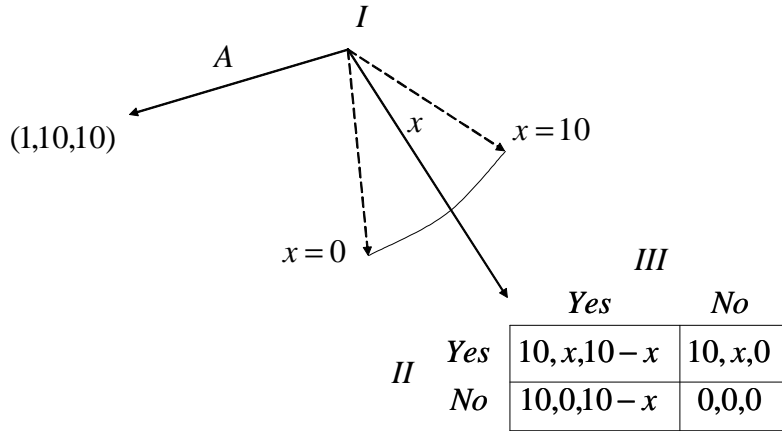


Figure 1: The extensive form game, Γ_1 . The choice x for player I ranges over the nonnegative integers, $\{0, 1, \dots, 10\}$.

The stage game for our first example is the extensive form Γ_1 , illustrated in Figure 1. We interpret the choice of $x \in \{0, 1, \dots, 10\}$ by player I as a bribe to player II (with $10 - x$ the bribe to player III). If player I chooses the cooperative action A , then the stage game ends. If player I chooses instead to bribe players II and III , players II and III simultaneously decide whether to accept or reject the bribe. Then, the stage game ends, and all actions become common knowledge. The game has many subgame-perfect equilibria, but they all share some common features: player I attempts to bribe the other players rather than behave cooperatively, and both players II and III accept any positive bribe offered. Moreover, the set of subgame perfect equilibrium payoffs is given by

$$\{(10, x, 10 - x) : x \in \{0, 1, \dots, 10\}\}.$$

Note that the payoffs of players II and III are negatively related across the equilibria of the stage game.

We are interested in the possibility of using these multiple equilibria to construct an equilibrium of the repeated game where player I behaves cooperatively in the first period.

We begin by arguing that it is impossible to support the cooperative play A by I in the first period using continuation play in the second period that is independent of the nature of I 's bribe: If player I deviates by attempting the bribe x , then equilibrium would require that players II and III both reject player I 's bribe. For II to reject, her continuation payoff from rejection must be at least x . If the continuation play is independent of I 's bribe, this implies that II 's continuation payoff after rejection is 10

(otherwise *II* would accept a bribe of 10). At the same time, for *III* to reject, his payoff must be at least $10 - x$, again for all x . But this requires that *III*'s continuation payoff after rejection is also 10, which is impossible.

On the other hand, it is easily verified that the following profile is a subgame-perfect equilibrium: In the first period, player *I* plays cooperatively, and all bribes are rejected. In the second period, after cooperative play in the first period, player *I* bribes at some level x in the second period (any level works). After a bribe of x is rejected by both *II* and *III* in the first period, player *I* offers the *same* bribe in the second period, which is accepted by both *II* and *III*. (If *I* offers another bribe in the second period, both players *II* and *III* accept.) If only one player accepts a bribe in the first period, then in the next period, player *I* offers the bribe that leaves the accepting player with 0. It is irrelevant whether that player accepts the bribe in the second period (the other player of course accepts). Finally, if both players accept the first-period bribe, an arbitrary continuation equilibrium is played (since both players accepting is a simultaneous deviation by *II* and *III*, these payoffs are irrelevant for the purposes of checking for subgame perfection).

While we assume here that the stage game is repeated only once, the same logic clearly holds if the stage game is infinitely repeated and payoffs are discounted. The play of *A* in every period of the infinitely repeated game can be sustained as an equilibrium outcome using a simple penal code if and only if the discount factor satisfies $\delta \geq 2/3$ (using the continuation of perpetual $x = 5$ and *Yes, Yes* if a bribe is accepted). However, there exists an equilibrium using the non-simple punishments described above sustaining the same outcome for $\delta \geq 1/2$.

This example shares some features with the three-person alternating-offer bargaining game of Shaked (described in Sutton (1986, p. 721)). When players are sufficiently patient, any division of the pie between the three bargainers is consistent with subgame perfection. Those equilibria, like here, specify different continuation equilibria as a function of the identity of the player who is supposed to reject (in his game, one veto is enough). Our example illustrates that the continuation equilibria may need to be finely tuned to the original deviation. In Shaked's example, it is sufficient to promise the entire pie to the rejecter, while both players must be rewarded in our game, since both must reject.

3. Second example: The punishment should fit the crime

“Is it her fault or mine?

The tempter or the tempted - who sins the most?”

William Shakespeare *Measure for Measure*, Act 2, Scene 2.

Our first example incorporates a conflict of interest between the players who are not supposed to acquiesce to a deviation. We now turn to our second example, which has

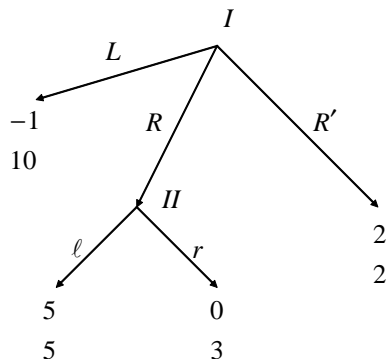


Figure 2: The extensive-form game Γ_2 .

	A	B
A	5, 5	0, 0
B	0, 0	1, 1

Figure 3: The coordination game

a strong coordination flavor and highlights the complications arising from a trade-off between punishing the deviator and rewarding the punishing player for applying the costly punishment.

We begin with the extensive-form game Γ_2 , presented in Figure 2. In this stage game, the unique backward induction equilibrium is for player I to play R , and player II to play ℓ after R . Player II of course, prefers that player I play L .

Before we describe the repeated extensive-form game, we first analyze a simpler two-period game. In the first period, Γ_2 is played, while in the second period, the coordination game of Figure 3 is played. Each player's payoff is the sum of his or her payoffs in the two periods. There is also a public random variable, ω , realized just before this coordination game is played. The random variable is independent of play, being uniformly distributed on $[0, 1]$; it serves only as a public correlating device.² By

²We can dispense with the assumption of a public correlating device by using jointly controlled lotteries, introduced by Aumann, Maschler, and Stearns (1968). Suppose the two players simultaneously announce a number from $[0, 1]$. Let ω equal their sum, if the sum is less than 1, and equal their sum minus 1 otherwise. It is easy to verify that if one player uniformly randomizes over his selection, then ω

using this device, second-period pure-strategy play can yield any payoff to the players between $(1, 1)$ and $(5, 5)$.³ In the second period, the payoffs of the two players are, by construction, identical. It is impossible to punish (or reward) one player without simultaneously doing the same to the other.

We are interested in equilibria that support L as a choice by player I in the first period. The payoffs in this example have been chosen so that the variation in second-period payoffs alone is insufficient to deter player I from playing R when player II chooses l , her myopic best reply: Player I 's first-period incentive to deviate by playing R is 6 (when II chooses l), while the largest punishment the second period can impose is less than 5. On the other hand, if player II could be induced to play r , we will see that for some specifications of second-period play, player I would not find R profitable.

Any equilibrium in which player I plays L rather than R must then provide incentives for player II to play r within the period. The coordination nature of the second-period game implies that these incentives must be fine-tuned. Maximizing player I 's payoff after L only relaxes his incentive constraints, and so we can assume AA is played after L . Similarly, we can suppose BB is played for sure after Rl , which yields a second-period payoff of 1 to each player. Let $1 \leq y(R, r) \leq 5$ denote the second-period payoff to each player after (R, r) . For player II to choose r rather than l after R , her payoff to r , which is 3, followed by $y(R, r)$ in the second period must exceed her payoff to l , which is 5, followed by 1 in the second period. That is, $3 + y(R, r) \geq 5 + 1$, or $y(R, r) \geq 3$. In the same way, for player I to prefer L over R , we must have $0 + y(R, r) \leq -1 + 5$, i.e., $y(R, r) \leq 4$.

Consider now player I 's incentive to deviate from L by choosing R' . Payoffs are such that this deviation is unprofitable if and only if it is followed by a severe enough punishment of player I in the second period. By choosing R' rather than L , player I increases his first-period payoff by 3, but his second-period payoff can be reduced from 5 to less than 2. Let $1 \leq y(R') \leq 5$ denote each player's second-period payoff after the play of R . For player I to prefer L over R' , we need $2 + y(R') \leq -1 + 5$, i.e., $y(R') \leq 2$.

Hence, equilibrium play of L requires that each player's second-period payoff is $3 \leq y(R, r) \leq 4$ after (R, r) , and $y(R') \leq 2$ after R' . That is, the equilibrium in the second-period subgame after (R, r) is necessarily different from that after R' .⁴ The play of L in the first period cannot therefore be sustained by a "simple" penal code in which continuation play is independent of the particular deviation chosen by player I . However, the play of L can be sustained by a more complex strategy profile which

is uniformly distributed on $[0, 1]$, for any choice by the other player. Consequently, it is an equilibrium for each player to uniformly mix.

³Any payoff $x \in [1, 5]$ for each player can be achieved by correlating between AA and BB , with probability $(x - 1)/4$ on AA . Allowing mixing yields payoffs as low as $(\frac{5}{6}, \frac{5}{6})$.

⁴For simplicity, our discussion focusses on pure strategies. As should be clear, neither the conclusion nor the logic depends on this restriction.

	<i>A</i>	<i>B</i>
<i>A</i>	Γ_2	0, 0
<i>B</i>	0, 0	1, 1

Figure 4: A representation of Γ_2^* .

	<i>Al</i>	<i>Ar</i>	<i>B</i>
<i>AL</i>	-1, 10	-1, 10	0, 0
<i>AR</i>	5, 5	0, 3	0, 0
<i>AR'</i>	2, 2	2, 2	0, 0
<i>B</i>	0, 0	0, 0	1, 1

Figure 5: The normal form of Γ_1^* .

employs different punishments after different deviations.

Consider now the extensive-form game, Γ_2^* , where the two players first simultaneously choose *A* or *B*, Γ_2 follows a simultaneous announcement of *A*, and the other payoffs are as described in Figure 4. Since Γ_2 has a unique subgame-perfect equilibrium with payoffs (5, 5), the coordination game of Figure 4 is identical to that of Figure 3.

The game Γ_2^* is played in two periods. As before, at the end of the first period, there is a public realization of a correlating random variable (see footnote 2).

The discussion above shows that there is a subgame-perfect equilibrium of the repeated game in which in the first period, both players choose *A*, and then player *I* chooses *L*. (The only issue we have not addressed is a deviation by *I* or *II* to *B* in the first period. Player *II* clearly cannot benefit from such a deviation. For player *I*, the period 1 payoff from this deviation is 0, the same as from *Rr* in Γ_2 , and the second-period play is the same as well, and so the deviation is not profitable.)

It is also an implication of the above discussion that this outcome cannot be achieved in any equilibrium with the property that play in the second period is independent of the nature of *I*'s deviation.

To conclude our discussion of this example, it is useful to compare our analysis with that of an analysis of the repeated normal form of Γ_2^* . The normal form is given in Figure 5. Treating the simultaneous-move normal form of Figure 5 as the stage game, the profile in which the profile (*AL*, *Ar*) is played in the first period and (*AR*, *Al*) in the second, with *any* deviation by player 1 resulting in *BB* in the second period, is a subgame-perfect equilibrium of the repeated game. However, this ignores the failure

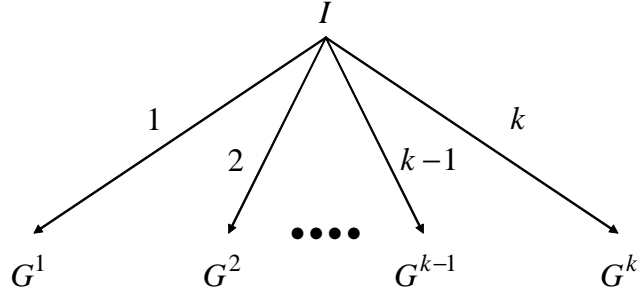


Figure 6: An extensive form game, Γ^* , with three players. Each G^ℓ is a two-player simultaneous move game between players *II* and *III*.

of player *II*'s decision to be optimal after R , an issue that cannot be addressed using subgame perfection in the repeated *normal* form. We return to this issue in Section 4.

4. Discussion

Our two examples illustrate how the logic of simple penal codes can fail in repeated extensive-form games because some equilibria require the use of within-period myopically-suboptimal “punishments” to ensure that deviations are not profitable. Figure 6 illustrates the simplest class of games in which simple penal codes fail. In the stage game Γ^* , player *I* moves first, choosing an action from the set $A \equiv \{1, 2, \dots, k-1, k\}$. Players *II* and *III*, after observing the action choice ℓ , play a finite simultaneous-move game $G^\ell \equiv \{(B_\ell, u_2), (C_\ell, u_3)\}$, where, for $i = 1, 2, 3$, $u_i(\ell, b, c)$ is player i 's payoff from the action profile (ℓ, b, c) , $\ell \in A$, $(b, c) \in B_\ell \times C_\ell$. Some of these games G^ℓ may be trivial, in that players *II* and *III* have only a single action. Observe that our two examples fall into this class. Denote the normal form of Γ^* by G^* .

The game in Figure 6, Γ^* , is played twice, with payoffs summed over the two periods. Suppose there is a subgame-perfect equilibrium, σ , with first-period action profile $(1, (b_\ell, c_\ell)_\ell)$, so that $(1, b_1, c_1)$ is played in the first period (in our examples, G^1 is trivial). If, for some $\ell \neq 1$, there is another Nash equilibrium of G^ℓ , $(\hat{b}_\ell, \hat{c}_\ell)$, giving player *I* a higher payoff than (b_ℓ, c_ℓ) (i.e., $u_I(\ell, b_\ell, c_\ell) < u_I(\ell, \hat{b}_\ell, \hat{c}_\ell)$), then it seems natural to interpret (b_ℓ, c_ℓ) as a *within-period punishment*, since it lowers *I*'s payoff from deviating to ℓ . Moreover, if (b_ℓ, c_ℓ) is itself a Nash equilibrium of G^ℓ , we say the within-period punishment is *myopically optimal*.

In a simple penal code for the repeated game of Γ^* , there is a *single* (second-period) continuation equilibrium of Γ^* that is used to punish player *I* for deviating from 1 in

the first period, and so the *same* equilibrium of Γ^* is specified after each $\ell \neq 1$. Our two examples show that there are games in which a subgame-perfect first-period action profile cannot be supported as first-period choices in a simple penal code.

Before we discuss the examples, we offer some general comments on simple penal codes in the context of the repeated Γ^* . The subgame perfect equilibrium σ specifies a continuation profile $\sigma|_\ell$ after each first-period profile (ℓ, b_ℓ, c_ℓ) (which are, of course, off-the-equilibrium path for $\ell \neq 1$). Let $\sigma|_{\hat{\ell}}$ be the (second-period) continuation profile that minimizes player I 's payoff in Γ^* . Consider now the modification of σ where we replace all continuation profiles by $\sigma|_{\hat{\ell}}$ in the second period after any choice $\ell \neq 1$ by player I and any choice by the other players. In other words, the new profile is a simple penal code with respect to player I . Clearly, this modified profile is a Nash equilibrium of the repeated Γ^* . More importantly for our purposes, this profile is also a subgame-perfect equilibrium of the repeated game *treating the second-period Γ^* 's as the only nontrivial subgames*. That is, we view the repeated game as a repetition of the normal form G^* . This is just the standard simple penal code result.

If (b_ℓ, c_ℓ) is a Nash equilibrium of G^ℓ for all $\ell \neq 1$ (so that all within-period punishments are myopically optimal), then this profile is clearly a subgame-perfect equilibrium of the repetition of Γ^* . However, if for some $\ell \neq 1$, (b_ℓ, c_ℓ) is not a Nash equilibrium of G^ℓ , the profile cannot be a subgame-perfect equilibrium of the repetition of Γ^* : at that ℓ , one of II or III has a myopic incentive to deviate, and they are not punished in the second period (since $\sigma|_{\hat{\ell}}$ is played regardless of their actions).

Moreover, it may not be possible to specify appropriate continuation profiles *independent of player 1's choice ℓ* so that the potentially deviating II or III have a sufficient incentive to play (b_ℓ, c_ℓ) . Two types of issues arise, highlighted by our two examples. In our second example, the payoffs of players I and II have a strong positive relationship (there is no player III), so that any continuation profile $\sigma|_\ell$ with a low payoff for player I also has a low player- II payoff. Thus, there is a conflict between rewarding player II for playing the within-period punishment strategy and imposing sufficient punishment on player I in the continuation game. This conflict must be carefully resolved by specifying play of the continuation equilibrium (A, A) , which is beneficial for player II , only after her punishment was particularly effective (i.e, after (R, r)), whereas (B, B) should be selected instead after (R, l) or R' . The punishment used in the continuation game had to 'fit the (particular) crime' committed by player I in the first period.

This discussion suggests that the problem highlighted is rather more general than the particular examples initially suggest. Our second example had the feature that it was feasible for player II to inflict within-period punishment on player I *only* if player I had chosen some particular deviation (R) , and that punishment was infeasible or ineffective after some other deviation (R') . The same underlying principle clearly applies in some cases where player II finds it myopically optimal to punish player I after some action(s)

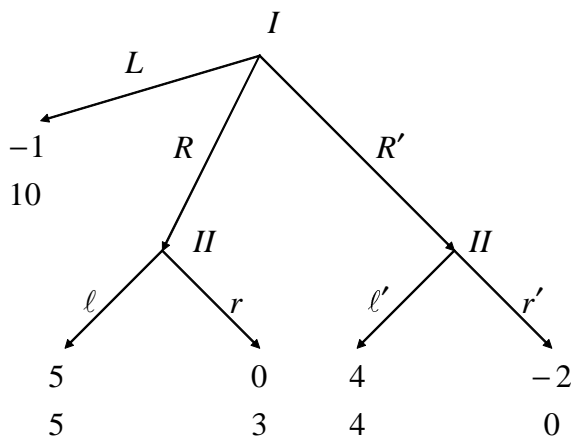


Figure 7: The extensive-form game Γ'_2 .

(i.e., R') but not others (i.e., R).⁵ More interestingly, the same type of considerations continue to apply when player II is required to inflict myopically-suboptimal punishment after all of player I 's deviations. A modified version of the extensive form Γ_2 , denoted Γ'_2 , is shown in Figure 7. Notice that in Γ'_2 , to support the candidate equilibrium play of L , player II must be induced to take the costly action r' after R' . If player II does not take this action, she can be punished by play of the worst continuation equilibrium, yielding 1 to each player. To be induced to take this action, therefore, player II must receive at least $1 + (4 - 0)$ in the continuation game, that is, the equilibrium yielding 5 to each player must be played after (R', r') ; this is sufficient to induce player I to play L . But note that if the same continuation equilibrium were played after (R, r) then player I would prefer to deviate to R ; after this history an equilibrium yielding no more than 4 (and no less than 3) to each player must be played.

In contrast to the coordination structure of the second example and its variants, the first example featured a conflict of interest between the later-moving players (II and III) who must both provide within-period punishment in order for player I 's deviation to be unprofitable. Here the problem lies not in the connection between player I and II 's 'incentive constraints', but in the connection between player II and III 's incentive constraints across *different* deviations by I . The deviating actions for player I involve offering some bribe x to player II (and $10 - x$ to player III), and $(b_\ell, c_\ell) \equiv (No, No)$ is a myopically-suboptimal within-period punishment for all $\ell \neq 1$. Here, the payoffs

⁵For example, after R' , give player II a choice between ℓ' and r' , with ℓ' myopically optimal and yielding payoffs (2, 2).

of *II* and *III* are inversely related, so that a continuation profile with a high payoff for player *II* necessarily has a low payoff for player *III*. At the same time, different bribes by player *I* imply different incentives for *II* and *III* to deviate from *No*. Consequently, the same continuation profile cannot simultaneously sufficiently reward both players for rejecting the bribe, independently of the size of the bribe. Rewarding one player more for within-period punishment means rewarding the other player less. Intuitively, to provide optimal incentives, the reward for a later-moving player should be positively related to the player's temptation not to punish the deviant player *I*. Since the size of this temptation depends on player *I*'s choice of action, it is necessary to fine-tune continuation play to the particular deviation chosen by player *I*. Thus when the interests of later-moving players are conflicting, not only should the punishment fit the crime, but the reward offered to punishing players should also be commensurate with the myopic temptation that they resisted.

5. Applications

The two examples we have discussed are intentionally simplified to clarify the issues and may seem highly special. Nevertheless, the same type of considerations highlighted arise naturally in many applications. The variant of the extensive-form game Γ_2 shown in Figure 7 can, for instance, be used to represent a game between an entrant firm (player *I*) and an incumbent (player *II*). The entrant may decide not to enter (play *L*), or if he enters, he can decide the "scale" of his entry: small (*R*) or large (*R'*). The incumbent must then decide whether to fight the entrant or not, which reduce both firms' profits and will be more costly to both firms when the entrant has chosen to enter on a large scale. This classic game has precisely the feature highlighted by the second example that the players' interests are ultimately aligned (they would prefer to avoid fierce competition) but the sequencing of moves is such that one player has the opportunity to punish the other for deviations within the period. Similar issues arise in games such as repeated multiple unit (e.g., English) auctions, sequential public good contribution games, or standard-setting when there are network externalities and firms prefer that their own technology is adopted as the industry standard.

As suggested by the discussion above, the extensive form Γ_1 in our first example has a natural interpretation as a bargaining game where player *I* has a pie of 10 to be split between himself and two others, with decisions being taken by majority voting (to make this interpretation literal, simply replace the payoff of 1 for player *I* after the action *A* with a payoff of zero). This type of game has many applications, for example, voting over budgetary allocations in a legislature or an election game where the politician (player *I*) must make concessions to obtain the support of one of two lobby groups (*II* and *III*) to be elected. Many natural applications fit an extended version of the game where player *I* can decide not only how his bribe should be allocated between players

II and *III* but also the total amount of the bribe. For example, player *I* may require a good which can be supplied by one or both of players *II* and *III* where the surplus from trade is 20.

A more extended application of the same type of difficulty in a very different setting is provided by Nocke and White (2003). That paper describes circumstances under which collusion can be sustained in an intermediate goods industry where several upstream firms compete to sell inputs to downstream retailers. In the stage game, which is infinitely repeated, the upstream firms first simultaneously make contract offers to downstream firms; then, the downstream firms simultaneously decide which contract(s) to accept; and, finally, the downstream firms compete in the retail market. An upstream firm can profitably deviate from the collusive equilibrium only if his deviant contract offer is accepted by at least one downstream retailer. Whilst Nocke and White concentrate on the simpler case where collusion is sustained by “Nash reversion”, it can be shown that the optimal punishment scheme induces downstream retailers to help sustain upstream collusion by rejecting certain deviant offers (in particular, those which are very profitable for the deviating firm). In order that downstream firms do indeed reject such offers, the tempted downstream retailers need to be “rewarded” by the play of an equilibrium in the continuation game that is favorable to them in that case. However, similar to our first example, the deviant upstream firm may “bribe” any individual downstream retailer to accept his deviant contract by sharing the rents from undercutting his rivals with the accepting downstream firm(s). Upstream collusion can be sustained for lower values of the discount factor if the reward to a each downstream retailer (i.e., the equilibrium in the continuation game) is tailored to the magnitude of the individual bribes, such that those downstream retailers receiving more tempting offers expect to receive more in future play.

6. Conclusion

In this paper we have shown by means of example that when players are impatient, non-simple penal codes, where the continuation play beginning in the period after a deviation depends on the particular deviation chosen at a given node, can be used to support a larger set of equilibria than would be supportable by the use of “simple” penal codes which are independent of the particular deviation chosen. Our examples suggest two particular features of games which seem to mitigate against the optimality of simple punishments: first, when within-period punishment should be inflicted by several later-moving players who have mutually-conflicting interests with regard to the continuation equilibrium; and second, when the interests of the (early-moving) deviator and the (later-moving) punisher are too closely aligned in the continuation game.

References

- ABREU, D. (1986): “Extremal Equilibria of Oligopolistic Supergames,” *Journal of Economic Theory*, 39(1), 191–225.
- (1988): “On the Theory of Infinitely Repeated Games with Discounting,” *Econometrica*, 56(2), 383–396.
- AUMANN, R. J., AND M. B. MASCHLER (1995): *Repeated Games with Incomplete Information*. MIT Press, Cambridge MA.
- AUMANN, R. J., M. B. MASCHLER, AND R. E. STEARNS (1968): “Repeated Games of Incomplete Information: An Approach to the Non-Zero Sum Case,” Report of the U.S. Arms Control and Disarmament Agency ST-143, reprinted in Aumann and Maschler (1995).
- FUDENBERG, D., AND E. MASKIN (1986): “The Folk Theorem in Repeated Games with Discounting or Incomplete Information,” *Econometrica*, 54, 533–554.
- NOCKE, V., AND L. WHITE (2003): “Do Vertical Mergers Facilitate Upstream Collusion?,” PIER Working Paper 03-033, University of Pennsylvania.
- RUBINSTEIN, A., AND A. WOLINSKY (1995): “Remarks on Infinitely Repeated Extensive-Form Games,” *Games and Economic Behavior*, 9(1), 110–115.
- SORIN, S. (1995): “A Note on Repeated Extensive Games,” *Games and Economic Behavior*, 9(1), 116–123.
- SUTTON, J. (1986): “Non-Cooperative Bargaining Theory: An Introduction,” *Review of Economic Studies*, 53(5), 709–724.