FACULTY OF LAW, ECONOMICS AND FINANCE

CREA Discussion Paper 2016-22

Economics

Centre for Research in Economics and Management
University of Luxembourg

Faking Patience in a Stochastic Prisoners' Dilemma

available online : http://wwwfr.uni.lu/recherche/fdef/crea/publications2/discussion_papers

Elena M. Parilina, Saint Petersburg State University, Russia
Alessandro Tampieri, University of Greenwich and CREA, University of Luxembourg

November, 2016

For editorial correspondence, please contact: crea@uni.lu
University of Luxembourg
Faculty of Law, Economics and Finance
162A, avenue de la Faïencerie
L-1511 Luxembourg

UNIVERSITÉ DU LUXEMBOURG

Faking Patience in a Stochastic Prisoners' Dilemma*

Elena M. Parilina† and Alessandro Tampieri‡

November 28, 2016

Abstract

This paper analyses tacit collusion in an infinitely repeated, stochastic prisoners' dilemma. We focus on a duopoly where two states of the world may occur and players adopt the grim-trigger strategy. Duopolists may be of three types, according to whether they choose a collusive behaviour in both states (patient), in none of the states (impatient) or in one state only (mildly patient). The presence of different states of the world affects the strategic role of beliefs. A mildly patient player has an incentive in "faking patience", which increases with the competitor's belief that the player is patient. Interestingly, this effect prevents collusive equilibria to occur when the belief of patience is strong.

JEL codes: C73, D43, L13.

Keywords: Repeated Games, Markov Perfect Bayesian Equilibrium.

*We are grateful to Nicolas Klein for helpful comments. The usual disclaimer applies.

†Saint Petersburg State University, 7/9 Universitetskaya nab., Saint Petersburg 199034, Russia; e.parilina@spbu.ru.

‡University of Greenwich and CREA, University of Luxembourg: tamp79@gmail.com.

1 Introduction

Tacit collusion emerges when firms collude without explicitly agree on that. Examples of these practices are limiting price cutting, excessive advertising or other forms of competition. Although not illegal, tacit collusion is a primary concern of antitrust authorities.

One of the ever interesting questions in industrial economics is whether and when firms choose collusion over competition. In this literature, collusion is modelled as a repeated oligopoly game. The folk theorem characterizes the set of equilibria and shows that collusion emerges if firms are patient enough (e.g., if the discount factor of future payoffs is sufficiently high). Friedman (1971) assumed perfect information and infinitely patient players. Fudenberg and Maskin (1986) and Radner (1986) assumed a discount factor asymptotic to one, and perfectly correlated information. Up to date folk theorems (Mailath and Samuelson 2006) support the idea that a collusive agreement may occur with limited information if firms are sufficiently patient. In a recent contribution, Harrington and Zhao (2012) examine tacit collusion in an infinitely repeated prisoners' dilemma with two types of players, patient and impatient, where a player's type is private information. If both players are patient and this information is common knowledge, then they can cooperate with a grim-trigger strategy. They find that the longer the players go without cooperating, the lower is the probability that they will cooperate in the next period.

So far, the analysis of collusion has abstracted from the presence of shocks, structural changes in the economy, or other external aspects that ultimately may affect the firms' decisions. A dynamic stochastic game (Shapley, 1953) is the right tool to catch these effects. The purpose of this paper is thus to investigate tacit collusion in a stochastic-game setting.

We examine a duopoly where two states of the world stochastically occur over time and players adopt a grim trigger strategy. The payoffs in the different states of the world are such that a not too patient player may collude in one state of the world and deviate in the other. This entails the presence of potentially 3 types of players: patient, adopting a cooperative behaviour in both states, impatient (never cooperative), and the mildly patient, who cooperates in one state ("good state") but deviates in the other ("bad state"). There are two phases in the repeated game. The first is the learning phase, where each player determines the competitor's types. The second is the collusion phase, where cooperation may emerge. In order to get tractable results, we focus the analysis in a short-term learning phase of one period.

The assumption of different states influences the strategic role of beliefs. Suppose that, in the learning phase, a patient player has a strong belief that the competitor is patient too. If a mildly patient competitor pretends to be patient and she is believed, a cooperative strategy will be played in the collusive phase; but then, this competitor will deviate in the bad state. Therefore, the mildly patient competitor has an incentive in "faking patience", which increases with the player's belief about the competitor's patience. In turn, since players rationally predict this type of behaviour, a strong belief in patience surprisingly will not lead to a fully cooperative equilibrium. In particular, the equilibrium strategy will exhibit cooperation in the good state and non-cooperation in the bad state of the world.

Many economic situations and interactions are well modelled as a stochastic game (Amir,2003). The analysis of stochastic games to problems emerging in industrial organisation has been developed in the contributions of Doraszelsky and Escobar (2016),Besanko *et al.* (2010), Pakes and McGuire (2001), Olley and Pakes (1996) and Bergemann and Välimäki (1996), *inter alia*. The present analysis combines the features of Harrington and Zhao (2012) who consider different types of players, with Rotemberg and Saloner (1986), Haltiwanger and Harrington (1991) and Bagwell and Staiger (1997), who evaluate the impact of business cycle on collusive behaviour.

The remainder of the paper is organised as follows. Section 2 introduces the model. Section 3 focusses the analysis on a class of equilibria for a two-phases game. Section 4 shows the results, while concluding remarks are in Section 5. All formal proofs can be found in the Appendix.

2 The model

2.1 Preliminaries

Consider a stochastic, infinitely-repeated Prisoners' Dilemma with two players. A player may be interpreted as a firm competing in a symmetric Cournot or Bertrand duopoly. Time is discrete and, in any time period $t = 1, 2, \ldots$, one of two states of nature $\{1, 2\}$ can be realised. States differ in their payoffs:

$$\begin{array}{cc} & \begin{array}{cc} C & D \end{array} \\ \begin{array}{c} C \\ D \end{array} & \begin{pmatrix} (a_s, a_s) & (c_s, b_s) \\ (b_s, c_s) & (d_s, d_s) \end{pmatrix}, \end{array} \quad s \in \{1, 2\}. \tag{1}$$

In each state for both players the first strategy C is "to cooperate" (to set a high, collusive price), the second strategy D is "to deviate" (to set a low, competitive price). Throughout the analysis, we will assume that deviation is more profitable in state 2 than in state 1.

The standard assumption of a prisoners' dilemma implies

$$b_s > a_s > d_s > c_s,$$

and

$$2a_s > b_s + c_s,$$

for any $s = 1, 2$. We denote the action set of any player i in any state s as $\mathbf{A} = \{C, D\}$ and the payoff function of any player as $u^s : \mathbf{A} \times \mathbf{A} \to R$, given by the payoff matrix (1) for state s.

We consider both possible initial states. In particular, we denote a player's payoff as a 2-dimensional vector where the first entry is the payoff in the game starting from state 1 and the second is the payoff for the game starting from state 2. The transition from a particular state does not depend on time period and the realised strategy profile. The probability that the game transits from state s to state 1 (and 2) is equal to π_s (and $1 - \pi_s$). A transition from a state to another one can be interpreted as a shock in the demand or supply of the industry.

Let δ be a discount factor of a player's payoff. We are interested in finding strategy profiles that are subgame perfect for some δ. The strategy profile is subgame perfect if, for any time period and any state, the corresponding part of the strategy profile is an equilibrium in the subgame. We assume that players adopt the grim trigger strategy, according to which cooperation occurs in equilibrium if the expected player's payoff in the whole game when she "cooperates" is higher than her expected player's payoff when she "deviates" at the current stage and then is punished by getting d_i in the rest of the game. Given this strategy, we are able to find the restrictions on the player's discount factor to define some strategy profiles (see below). Note that collusion may or may not be implemented according to the discount factor of both players.

2.2 Strategy profiles

We denote the strategy profile $\sigma = (\sigma_1, \sigma_2)$, where $\sigma_i = \{\sigma_i^s(t)\}_{t=1}^{\infty}$ such that $\sigma_i^s(t+1) : H(t) \to \mathbf{A} \times \mathbf{A}$ is an action of player i in time period $t+1$ when state s is realised,

$s \in \{1,2\}$, and $H(t) = ((s(1), \mathbf{a}(1)), \ldots, (s(t), \mathbf{a}(t)))$ is a history of time period t containing the sequence of realised states and corresponding action profiles $(s(t), \mathbf{a}(t))$, where $s(t)$ is the state in time period t and $\mathbf{a}(t)$ is the realised action profile in this state.

We are now in a position to consider the types of strategy profiles and the restrictions in δ such that they are subgame perfect equilibria. Begin by the definition of three strategy profiles.

Definition 1 *A "non-cooperative strategy profile" is denoted as* $\sigma_n = (\sigma_{n,1}, \sigma_{n,2})$, *where* $\sigma_{n,i} = \{\sigma_{n,i}^{s,t}\}_{t=1}^{\infty}$ *is such that* $\sigma_{n,i}^{s,t+1}(H(t)) = D$ *for any* $s = 1,2$, $t = 1,2,..$ *and any* $H(t) = ((s(1), \mathbf{a}(1)), \ldots, (s(t), \mathbf{a}(t)))$.

The non-cooperative strategy profile prescribes players to choose action D in any state in any time period, so that a Nash equilibrium occurs.

Definition 2 *A "cooperative strategy profile" is denoted as* $\sigma_c = (\sigma_{c,1}, \sigma_{c,2})$, *where* $\sigma_{c,i} = \{\sigma_{c,i}^{s,t}\}_{t=1}^{\infty}$ *is such that*

$$\sigma_{c,i}^{s,t+1}(H(t)) = \begin{cases} C, & \text{if } H(t) = H_c(t) \\ D, & \text{if there exists } t' \leqslant t \text{ such that} \\ & \mathbf{a}(t') = (C,D) \text{ or } (D,C) \\ \text{any from } \mathbf{A}, & \text{otherwise} \end{cases},$$

where $H_c(t) = ((s(1), (C,C)), \ldots, (s(t), (C,C)))$ *is a history at time period* t *where, in all previous time periods, both players choose strategy* C.

The cooperative strategy profile prescribes players to choose action C in this time period if the history shows past cooperation (i.e. no deviations are observed in the previous time periods). If players observe deviations from action profile (C,C) in the history, then they choose strategy D and realise the Nash equilibrium action profile (D,D).

Definition 3 *A "semi-cooperative strategy profile" is denoted as* $\sigma_{sc} = (\sigma_{sc,1}, \sigma_{sc,2})$, *where*

$\sigma_{sc,i} = \{\sigma^{s,t}_{sc,i}\}^{\infty}_{t=1}$ *is such that*

$$\sigma^{s,t+1}_{sc,i}(H(t)) = \begin{cases} C, & \text{if } s=1 \text{ and } H(t)=H_{sc}(t) \\ D, & \text{if } \{s=2 \text{ and } H(t)=H_{sc}(t)\} \text{ or} \\ & \text{there exists } t' \leqslant t \text{ such that} \\ & \mathbf{a}(t')=(C,D) \text{ or } (D,C) \\ \text{any from } \mathbf{A}, & \text{otherwise} \end{cases},$$

while $H_{sc}(t)$ *is a history in which all elements are of two types:* $(1,(C,C))$ *and* $(2,(D,D))$.

With a semi-cooperative strategy profile, players choose action C in state 1 and action D in state 2.

To complete the analysis, we introduce a final strategy profile.

Definition 4 *A "deviating strategy profile" is denoted as* $\sigma_d = (\sigma_{d,i}, \sigma_{c,j})$, *and* $\sigma_{d,i} = \{\sigma^{s,t}_{d,i}\}^{\infty}_{t=1}$ *is such that*

$$\sigma^{s,t+1}_{d,i}(H(t)) = \begin{cases} C, & \text{if } s=1 \text{ and } H(t)=H_c(t) \\ D, & \text{if } s=1 \text{ and } H(t)\neq H_c(t) \\ D, & \text{if } s=2. \end{cases} \quad (2)$$

In this profile one player j plays strategy $\sigma_{c,j}$ given by Definition 2 while player i realises strategy 2. This profile may occur when player j has a belief that the competitor i will play cooperatively while player i will in fact deviate in state $s=2$. In turn, when player j observes a deviation from cooperative strategy profile, she reacts with D in all stages afterwards according to strategy $\sigma_{c,j}$.

2.3 Discounted payoffs

In this section we calculate the players' payoffs discounted by δ. In what follows, we omit subscript i in the notation of the payoff. A player's discounted payoff when a strategy profile σ is realised is

$$V(\sigma) = \sum_{t=1}^{\infty} \delta^{t-1}\Pi^{t-1}U(t),$$

where

$$\Pi = \begin{pmatrix} \pi_1 & 1-\pi_1 \\ \pi_2 & 1-\pi_2 \end{pmatrix},$$

and $U(t) = (u^1(t), u^2(t))'$, $u^s(t) = u^s(t, a_1(t), a_2(t))$ is the payoff of the player in time period t and state s, corresponding to the strategy profile σ.

We denote the discounted payoffs of a player when players play non-cooperatively (choose action D in any state), cooperatively (choose action C in any state) and semi-cooperatively (choose action C in state 1 and action D in state 2) as V_n, V_c and V_{sc}, respectively, where subscripts n, c and sc stand for "non cooperative" "cooperative" and "semi-cooperative". First, we define $p_s = (\pi_s, 1-\pi_s)$ and

$$\widetilde{\Pi} = \frac{1}{(1-\delta)(1-\delta(\pi_1-\pi_2))} \begin{pmatrix} 1-\delta(1-\pi_2) & \delta(1-\pi_1) \\ \delta\pi_2 & 1-\delta\pi_1 \end{pmatrix}.$$

Lemma 1 *1. The discounted payoff of a player when players play non-cooperatively is*

$$V_n = \begin{pmatrix} V_n^1 \\ V_n^2 \end{pmatrix} = \widetilde{\Pi} \begin{pmatrix} d_1 \\ d_2 \end{pmatrix}, \tag{3}$$

where V_n^s is a payoff of the player starting from state $s = 1, 2$.

2. The discounted payoff of a player when players play cooperatively is

$$V_c = \begin{pmatrix} V_c^1 \\ V_c^2 \end{pmatrix} = \widetilde{\Pi} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}. \tag{4}$$

3. The discounted payoff of a player when players play semi-cooperatively is

$$V_{sc} = \begin{pmatrix} V_{sc}^1 \\ V_{sc}^2 \end{pmatrix} = \widetilde{\Pi} \begin{pmatrix} a_1 \\ d_2 \end{pmatrix}, \tag{5}$$

4. Let player i realise strategy $\sigma_{c,i}$ and player j realise strategy $\sigma_{d,j}$. The discounted payoff of player j (the deviating player) is

$$V_d = \begin{pmatrix} \frac{1}{1-\delta\pi_1}[a_1 + \delta(1-\pi_1)(b_2 + \delta p_2 V_n)] \\ b_2 + \delta p_2 V_n \end{pmatrix}. \tag{6}$$

Proof. See Appendix. ∎

We are now in a position to define, for every state s, the critical value of δ_s for which each strategy profile is a subgame perfect equilibrium. For convenience, define

$$\delta_1^* \equiv \frac{b_1 - a_1}{p_1(V_{sc}(\delta_1^*) - V_n(\delta_1^*))}, \quad \delta_2^* \equiv \frac{b_2 - a_2}{p_2(V_c(\delta_2^*) - V_n(\delta_2^*))}, \tag{7}$$

where $\delta_1^* \in (0, 1)$ and $\delta_2^* \in (0, 1)$. Without loss of generality, we assume that, in state 1, deviation is less profitable than in state 2.

Assumption 1 *Let* $\delta_2^* > \delta_1^*$.

Assumption 1 allows us to consider a discounted payoff and a strategy profile in which cooperation is obtained in state 1 only. Proposition 1 summarises the conditions on the discount factor that determine which strategy profile is a subgame perfect equilibrium.

Proposition 1 *A cooperative strategy profile is subgame perfect equilibrium for any* $\delta \geqslant \delta_2^*$*. A semi-cooperative strategy profile is subgame perfect equilibrium if* $\delta \geqslant \delta_1^*$*. A non-cooperative strategy profile is always a subgame perfect equilibrium for any* $\delta \in (0, 1)$*.*

Proof. See the appendix. ∎

2.4 Player's types

Throughout the paper, we will make the following assumption on players' types.

Assumption 2 *A player can be one of three types:*

i. I *(impatient), whose discount factor* δ_I *satisfies* $\delta_I < \delta_1^*$*;*

ii. M *(mildly patient), whose discount factor* δ_M *satisfies* $\delta_1^* \leqslant \delta_M < \delta_2^*$*;*

iii. P *(patient), whose discount factor* δ_P *satisfies* $\delta_P \geqslant \delta_2^*$*.*

The cooperative strategy profile can be realised only if both players have type P. A semi-cooperative strategy profile can be realised if (i) both players are of type P, (ii) both players are of type M or (iii) if one is of type P and the other is of type M. If at least one of two players is of type I, then neither a cooperative nor a semi-cooperative strategy profile is played. If at least one player has discount factor less than δ_1^*, then only a non-cooperative strategy profile can be realised.

3 Two-phases game

Possibly, this game may exhibit several types of equilibria. Following Harrington and Zhao (2012), we analyse a class of equilibria for a two-phases game.

The first phase is *learning*, where the players' discount factors are private information and players try to recognise the competitor's type. This implies that players' strategies focus on beliefs of the other player's type, and not on the game history. The maximal duration of the learning phase is a finite T. It starts from time period $t = 1$ and may end before T in the case when players' types become common information or one of the players is identified as type I.[1] Cooperation cannot be supported by grim trigger strategies, and players are assumed to play the non-cooperative strategy profile in the next phase of the game.

The second phase is *collusion*, where the strategy depends on the history of the game, and players realise one of the strategy profiles σ_n, σ_c or σ_{sc}. The strategy in the collusion phase is a mapping of the state and the information about players' types which players have had before the second phase starts. Therefore, the strategy profile in the collusion phase is a mapping from information they have (probabilities that the other player is of type P, M and I) and the state (1 or 2) to the set of behaviour strategy profiles $\{\sigma_c, \sigma_{sc}, \sigma_n\}$.

The player's strategy in the whole game consists of the strategy in the learning and in the collusion phase. The equilibrium concept is Partially Markov Perfect Bayesian Equilibrium (PMPBE, Harrington and Zhao, 2012). The Markov property is that the strategy in the learning phase in any time period t depends only on the beliefs on the competitor's type, and does not depend on the time period and on the history. However, it is "partial" because, in the collusion phase, the equilibrium is not a Markov Perfect Bayesian Equilibrium as players engage in behaviour strategies once the learning phase ends up. Indeed, in the collusion phase there is no updating information about a player's type.

3.1 Learning phase: rules for updating beliefs

In this section we describe how the process of learning the competitor's type takes place. Before time period t and state s are realised, a player believes the other player to be of type P with probability α_t^s, to be of type M with probability β_t^s, and to be of type I with

[1] Indeed, in this case there is no reason to try to cooperate.

probability $\gamma_t^s = 1 - \alpha_t^s - \beta_t^s$, for every $t = 1, \ldots, T$, $s \in \{1, 2\}$. Of course, if player chooses strategy C in state 2, she has revealed her type as P, because she is the only type to play C in state 2. Conversely, if a player chooses strategy C in state 1, she may be identified as a P or M type. We assume that beliefs are public information and symmetric.

If state s is realised at period t, then for the next period $t+1$ the belief that the competitor is a P type is

$$\alpha_{t+1} = \begin{cases} \alpha_{t+1}^1, & \textit{for } s = 1 \textit{ in } t \\ \alpha_{t+1}^2, & \textit{for } s = 2 \textit{ in } t \end{cases}.$$

The same rule applies to β_{t+1} and γ_{t+1}. We assume that the initial beliefs about the other player's type are given, i.e. $\alpha_1 \in (0,1)$, $\beta_1 \in (0,1)$ are known and $\alpha_1 + \beta_1 \in (0,1)$.

The players' strategies in the learning phase depending of their types are as follows. I chooses D in any time period and any state; P chooses C in state $s \in \{1, 2\}$ in time period t with probability q_t^s; M chooses action C in state 1 in time period t with probability r_t^1, and action D in state 2 in any time period with probability 1.

The Markovian strategy of a P player in the learning phase is $q_t^s(\cdot) : [0,1] \times [0,1] \to [0,1]$ which is a probability of choosing action C in state s in time period t. It is a function of α_t and β_t of the other player to be of type P or M, respectively, $t = 1, \ldots, T$ and $s = 1, 2$. The strategy of an M player is $r_t^1(\cdot) : [0,1] \times [0,1] \to [0,1]$ which is a probability of choosing action C in state 1 in time period t. It is also a function of α_t and β_t, $t = 1, \ldots, T$. Finally, notice that a type M player in state $s = 2$ never chooses C. We denote the set of strategies of a type P player by Q, and of a type M player by R.

We use the Bayes rule to update probabilities α_t^s and β_t^s over time, $s \in \{1, 2\}$. First, consider the updating rule for state $s = 1$. If a player chooses C in time t, she is identified as type[2]

$$\begin{cases} I \quad \text{with prob } \gamma_{t+1}^s = 0; \\ P \quad \text{with prob } \alpha_{t+1}^1 = \dfrac{\alpha_t q_t^1}{\alpha_t q_t^1 + \beta_t r_t^1}; \\ M \quad \text{with prob } \beta_{t+1}^1 = 1 - \alpha_{t+1}^1 = \dfrac{\beta_t r_t}{\alpha_t q_t^1 + \beta_t r_t^1}. \end{cases} \tag{8}$$

[2]The probabilities α_{t+1}^1 and β_{t+1}^1 are defined if $\alpha_t q_t^1 + \beta_t r_t \neq 0$. In the case when $\alpha_t q_t^1 + \beta_t r_t^1 = 0$, it is impossible to observe action C in state $s = 1$.

If a player chooses D in state $s = 1$, she is identified as type[3]

$$\begin{cases} I \quad \text{with prob } \gamma_{t+1}^{s} = 1 - \alpha_{t+1}^{1} - \beta_{t+1}^{1} = \dfrac{1-\alpha_t-\beta_t}{1-\alpha_t q_t^1 - \beta_t r_t^1}; \\ P \quad \text{with prob } \alpha_{t+1}^{1} = \dfrac{\alpha_t(1-q_t^1)}{1-\alpha_t q_t^1 - \beta_t r_t^1}; \\ M \quad \text{with prob } \beta_{t+1}^{1} = 1 - \alpha_{t+1}^{1} = \dfrac{\beta_t(1-r_t)}{1-\alpha_t q_t^1 - \beta_t r_t^1}. \end{cases} \tag{9}$$

Consider next the updating rule for state 2. As already discussed, if a player chooses C is identified as type P with probability 1. On the other hand, if a player chooses D, she is identified as type[4]

$$\begin{cases} I \quad \text{with prob } \gamma_{t+1}^{2} = 1 - \alpha_{t+1}^{2} - \beta_{t+1}^{2} = \dfrac{1-\alpha_t-\beta_t}{1-\alpha_t q_t^2}; \\ P \quad \text{with prob } \alpha_{t+1}^{2} = \dfrac{\alpha_t(1-q_t^2)}{1-\alpha_t q_t^2}; \\ M \quad \text{with prob } \beta_{t+1}^{1} = 1 - \alpha_{t+1}^{1} = \dfrac{\beta_t(1-r_t)}{1-\alpha_t q_t^1 - \beta_t r_t^1}. \end{cases} \tag{10}$$

If, in time period $t-1$, state s has been realised, then the player updates the beliefs about the other player's type and calculates α_t^s, β_t^s, γ_t^s which will be the beliefs for time period t. To ease notation, we omit subscript s in the the probabilities α_t, β_t, and γ_t.

3.2 Collusion phase

The learning phase may end in time $t^* \leqslant T$, so that the collusion phase starts from $t^* + 1$. The rule to determine a behaviour strategy profile σ for the collusion phase is a function of probabilities α_{t^*+1} and β_{t^*+1}. To give a strict definition of equilibrium, suppose for simplicity that the learning phase lasts T periods. At the end of the learning phase, the beliefs that the competitor is of type P or M are α_{T+1} and β_{T+1}, respectively. Thus in the collusion phase, the strategy profile is $\sigma(\alpha_{T+1}, \beta_{T+1})$.

We suppose that players may realise one of three possible profiles in the collusion phase:

[3] The probabilities α_{t+1}^1, β_{t+1}^1 and γ_{t+1}^1 are defined if $\alpha_t q_t^1 + \beta_t r_t^1 \neq 1$. In the case when $\alpha_t q_t^1 + \beta_t r_t^1 = 1$, it is impossible to observe action D in state $s = 1$.

[4] The probabilities α_{t+1}^1, β_{t+1}^1 and γ_{t+1}^1 are defined if $\alpha_t q_t \neq 1$. In the case when $\alpha_t q_t^2 = 1$ ($\alpha_t = q_t^2 = 1$), it is impossible to observe action D in state $s = 2$.

$\sigma(\alpha_{T+1}, \beta_{T+1}) \in \{\sigma_c, \sigma_{sc}, \sigma_n\}$. The rule that determines the strategy profile chosen is

$$\sigma(\alpha_{T+1}, \beta_{T+1}) = \begin{cases} \sigma_c, & \text{if } \alpha_{T+1} = 1, \\ \sigma_{sc}, & \text{if } \alpha_{T+1} + \beta_{T+1} = 1 \text{ and } \alpha_{T+1} \neq 1, \\ \sigma_n, & \text{otherwise.} \end{cases} \tag{11}$$

We assume that the two players use the symmetric behaviour strategies, i.e., each player adopts the same rule to determine her strategy in the collusion phase. The payoff of a type i player, $i \in \{P, M, I\}$ is the sum of her payoffs in two phases of the game:

$$\Phi_i(q, r) = \sum_{t=1}^{T} \delta^{t-1} \Pi^{t-1} U_i(q_t, r_t) + \delta^T \Pi^T V(\sigma(\alpha_{t+1}, \beta_{t+1})), \tag{12}$$

where $q = \{q_t^s\}_{t=1,\dots,T, s=1,2}$ is a strategy of a type P player, $r = \{r_t^1\}_{t=1,\dots,T}$ is a strategy of a type M player, and

$$V(\sigma(\alpha_{t+1}, \beta_{t+1})) = \begin{cases} V_c, & \text{if } \sigma(\alpha_{t+1}, \beta_{t+1}) = \sigma_c \\ V_{sc}, & \text{if } \sigma(\alpha_{t+1}, \beta_{t+1}) = \sigma_{sc} \\ V_n, & \text{if } \sigma(\alpha_{t+1}, \beta_{t+1}) = \sigma_n \end{cases}.$$

The first part in the RHS of (12) is the payoff during the learning phase, while the second part is the payoff during the collusion phase. Note that the payoff in the learning phase depends on the player's type, but in the collusion phase the stage payoffs of the players of different types are equal, according to payoff matrix (1).

For convenience, denote the following strategy profiles.

Definition 5 *The strategy:*

- $(q^*||q, r^*)$ *is a strategy profile when a player of type P deviates from q^* and plays $q \in Q$, while the competitor implements strategy q^* if she is of type P and r^* if she is of type M;*
- $(q^*, r^*||r)$ *is a strategy profile when a player of type M deviates and plays $r \in R$ and the competitor implements strategy q^* if she is of type P and r^* if she is of type M.*

We are now able to define the Partially Markov Perfect Bayesian Equilibrium.

Definition 6 *A strategy profile (q^*, r^*) is a Partially Markov Perfect Bayesian Equilibrium (PMPBE) if for all players of type P and M and any $q \in Q$ and $r \in R$,*

$$\Phi_P(q^*, r^*) \geqslant \Phi_P(q^*||q, r^*),$$
$$\Phi_M(q^*, r^*) \geqslant \Phi_M(q^*, r^*||r).$$

4 Two-phase game with one-period learning phase

In this section we examine the case where the maximal length of the learning phase T is one time period. The fact that learning is arbitrarily assumed to terminate in some period is necessary in order to keep the analysis tractable.[5] We find the conditions when Partially Markov Perfect Bayesian Equilibrium exists in pure or mixed strategies of the learning phase.

4.1 Collusive equilibria

In this section we are interested in the equilibria where, in the collusion phase, the cooperative or semi-cooperative strategy profile is realised. These equilibrium strategies lead to cooperation (or semi-cooperation) in the collusion phase. The rule of choosing the strategy profile in the collusion phase is given by (11).

Since the player's payoffs V_n, V_{sc} and V_c are the function of discount factor δ and players have different discount factors, we update the notations of the players' payoffs by $V_n(\delta)$, $V_{sc}(\delta)$ and $V_c(\delta)$. Equations (3), (4) and (5) are the formulas calculating them. For convenience, we denote

$$\begin{aligned}
A_1 &\equiv \frac{d_1 - c_1}{d_1 - c_1 + a_1 - b_1 + \delta_P p_1(V_c(\delta_P) - V_n(\delta_P))}, \\
A_2 &\equiv \frac{d_1 - c_1}{d_1 - c_1 + a_1 - b_1 + \delta_M p_1\left(V_d(\delta_M) - V_n(\delta_M)\right)}, \\
A_3 &\equiv \frac{d_1 - c_1}{d_1 - c_1 + a_1 - b_1 + \delta_M p_1(V_{sc}(\delta_M) - V_n(\delta_M))}, \\
A_4 &\equiv \frac{d_2 - c_2}{d_2 - c_2 + a_2 - b_2 + \delta_P p_2(V_c(\delta_P) - V_n(\delta_P))}.
\end{aligned}$$

[5] Upon request, analytical computations of the problem with a 2-periods learning phase can be provided.

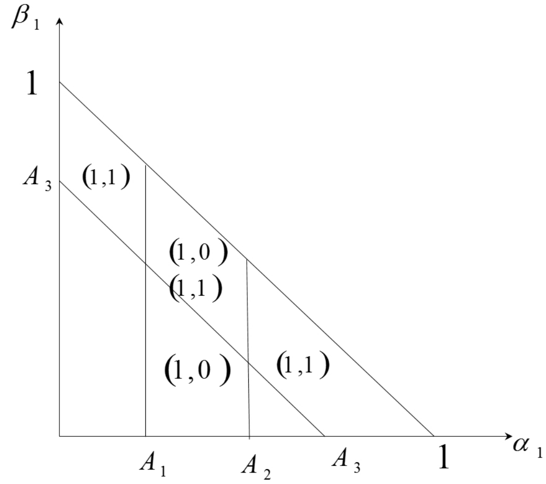

Figure 1: Equibria with initial state $s = 1$.

The next proposition summarises the equilibria that lead to collusion.

Proposition 2 *Let Assumption 2 hold. Consider a two-phase game and suppose* $T = 1$*. Then the collusive equilibria are the following:*

1. *when the initial state is* $s = 1$*:*

   1.i $(q_1^1, r_1^1) = (1, 0)$ *is a PMPBE if* $\alpha_1 \in [A_1; A_2]$.

   1.ii $\left(q_1^1, r_1^1\right) = (1, 1)$ *is a PMPBE if* $\alpha_1 + \beta_1 \geqslant A_3$;

2. *when the initial state is* $s = 2$*:*

   2.i $\left(q_1^2\right) = 1$ *is a PMPBE if* $\alpha_1 \geqslant A_4$.

   2.ii $\left(q_1^2\right) = q^*$ *is a PMPBE if* $\alpha_1 \geqslant A_4$*, with* $q^* = A_4/\alpha_1$.

Proof. See the appendix. ■

Figures 1 and 2 depict the regions of cooperative equilibria for initial state 1 and 2, respectively, in the space of the beliefs (α_s, β_s). The rule of updating beliefs in (8) and (9) helps to understand the strategy profile in the collusion phase. Suppose, for instance, that the game starts with state 1 and profile $(q_1^1, r_1^1) = (1, 0)$ is realised. If action C is observed, the updates beliefs are $a_2 = 1$, $\beta_2 = \gamma_2 = 0$, thus it is possible to recognise the other

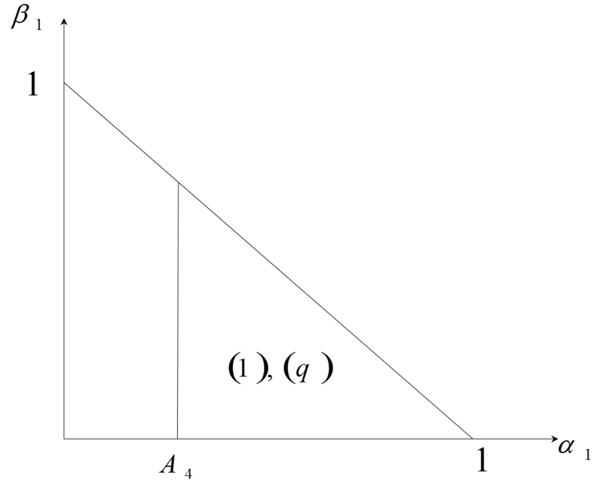

Figure 2: Equilibria with initial state $s = 2$.

player's type if she is of type P. Hence, if two players of type P meet, the equilibrium $(q_1^1, r_1^1) = (1, 0)$ leads to the cooperative strategy profile σ_c in the collusion phase. If even one of the two players is not of type P, the non-cooperative strategy profile σ_n will be realised in the collusion phase. Indeed, since the mildly patient and the impatient type adopt the same strategy, a patient player cannot recognise from the learning phase if the competitor is a mildly patient one, thus the semi-cooperative strategy is never played in the collusion phase.

When the game starts with state 1 and profile $(q_1^1, r_1^1) = (1, 1)$ is realised, i.e., players of types M and P cooperate with probability 1, the beliefs of a competitor's type after observing C are:

$$\alpha_2 = \frac{\alpha_1}{\alpha_1 + \beta_1}, \beta_2 = \frac{\beta_1}{\alpha_1 + \beta_2}, \gamma_2 = 0.$$

In this case there are positive probabilities that the competitor is either P or M. Hence the equilibrium strategy during the second phase is semi-cooperation σ_{sc}, which allows collusion in future states 1 and deviation in future states 2. This result emerges as a player does not recognise whether the competitor is of type P or M.

When the game starts with state 2, a type M cannot be identified, since she always deviates. Hence the belief β does not play any role in determining the equilibrium. However, a type P competitor is identified for sure. hence 2 P players realise the cooperative strategy profile σ_c in the collusion phase.

The next corollary compares the equilibrium payoffs in the parameter ranges where multiple equilibria occur, as a possible refinement in the equilibrium choice.

Corollary 1 *Suppose the game starts from state $s = 1$, and α_1, β_1 satisfy the conditions: $\alpha_1 \in [A_1, A_2]$ and $\alpha_1 + \beta_1 \in [A_3, 1]$. Then the payoff of type M player in equilibrium $(q_1^1, r_1^1) = (1, 1)$ is always not less than his payoff in equilibrium $(q_1^1, r_1^1) = (1, 0)$.*

Suppose the game starts from state $s = 2$, and $\alpha_1 \geqslant A_4$, then the payoffs of type P players in equilibrium $q_1^2 = 1$ is not less than their payoffs in equilibrium $q_1^2 = q^*$.

Proof. See the appendix. ■

Corollary 1 suggests equilibrium $(1, 1)$ as a refinement of multiple equilibria in state 1, and the pure strategy over the mixed one in state 2. In particular for state 1, this result intuitively suggests that, when the beliefs that the competitor is P or M are similar, it is unlikely to reach a result of full collusion. Indeed, the outcome is a semi-cooperative strategy in the collusion phase, i.e., the collusive strategy associated to the profile $(1, 1)$.

The next proposition summarises some comparative statics on the equilibrium payoffs with respect to belief.

Proposition 3 *The equilibrium payoffs of the players of types P and M are increasing functions of belief α_1. The payoffs of the players of types P and M in equilibrium $(q_1^1, r_1^1) = (1, 1)$ are the increasing functions of belief β_1.*

Proof. See the appendix. ■

Proposition 3, together with Proposition 2, state a surprising result: a strong belief that the competitor is of type P does not lead to a collusive equilibrium in the collusion phase. This is immediately evident by looking at Figure 1. The reason is simple. A very high α_1 gives a strong incentive to an M type to fake patience, that is, to act as a P type to lure the competitor to play a collusive strategy in the collusion phase. Indeed, if a collusive strategy is played in the collusion phase and state 2 occurs at some period, then the M-type player would deviate from collusion, thus tricking a P-type competitor. Given that players are aware of the "faking patience" effect, when the competitors are believed to have high patience, a semi-collusive equilibrium occurs.

4.2 Non-collusive outcome

We now turn to the PMPBE which leads to the non-collusive results, i.e., in the collusion phase of the game, the non-cooperative strategy profile σ_n is realised. The results are summarised in the following proposition.

Proposition 4 *Let Assumption 2 hold. Consider a two-phase game and suppose $T = 1$. Then the strategy profiles $(q_1^1, r_1^1) = (0, 0)$ and $(q_1^2) = (0)$ are non-collusive PMPBE for the game with the initial state $s = 1$ and $s = 2$ respectively.*

Proof. See the appendix. ■

The result in Proposition 4 shows the non-collusive behaviour when the learning phase lasts one-period, but this result can be generalised. In particular, we are able to show that, if $T > 1$ and the $t = 1$ strategy is $q_1^s = r_1^1 = 0$, $s = 1, 2$, then players will use the same profiles in any time period during the learning phase. This result was proved for the one-state game in Harrington and Zhao (2012), and holds for the two-states game. To prove this we consider the game where players of types P and M play action D in any state of time period 1, i.e., they realise the strategy profile $\{(q_t^s, r_t^1)\}_{t=1,\dots,T,s=1,2}$, where $q_1^s = r_1^1 = 0$, $s = 1, 2$. Hence any player observes only action D in any state of period 1 and updates her beliefs accordingly. In each state of period 2, the beliefs remain the same as in period $t = 1$. This easily follows from the assumption $q_1^s = r_1^1 = 0$, $s = 1, 2$. Using the Markovian property of the strategies, the players' strategies remain the same because they depend on the beliefs (which do not change) and not on the history. Therefore, the strategies of time period 2 are $q_2^s = q_1^s = 0$, $r_2^1 = r_1^1 = 0$. By induction, the same occurs in period $t = 3, \dots, T$ of the learning phase. Thus the strategy profile is $(0, 0)$ for state $s = 1$ and (0) for state $s = 2$. In turn, during the collusion phase the non-cooperative strategy profile σ_n will be realised.

5 Concluding remarks

In this paper we have analysed tacit collusion in a stochastic, infinitely repeated prisoners' dilemma. The starting point of our study is the deterministic prisoners' dilemma of Harrington and Zhao (2012), which is extended in a stochastic form with two states of the world.

We have shown that the presence of different states of the world drastically affects the strategic role of beliefs. A competitor that shifts from collusion to deviation according to the state of the world has an incentive in faking patience in the good state. Since this behaviour is expected and increases with the belief of patience, the latter loses its role in choosing collusion.

Our results have been outlined by focussing on a short-period learning phase. Albeit this last restriction somewhat limits the generality of our results, it keeps the analysis sufficiently tractable to highlight the emergence of the faking patience effect in the presence of stochastic states of the world.

References

[1] Amir, R. 2003. Stochastic games in economics: The lattice-theoretic approach. In A. Neyman and S. Sorin (eds.) *Stochastic Games and Applications*, NATO Science Series C, Mathematical and Physical Sciences **570**: 443-453.

[2] Bagwell, K. and Staiger, R. 1997. Collusion over the business cycle. *RAND Journal of Economics* **28**: 82-106.

[3] Bergemann, D. and Välimäki, J. 1996. Learning and strategic pricing, *Econometrica* 64: 1125–1149.

[4] Besanko, Y.D., Doraszelsky, U., Kryukov, Y., and Satterthwaite, M. 2010. Learning by ding, organizational forgetting and industry dynamics. *Econometrica* **78**: 453-508.

[5] Doraszelsky, U. and Escobar,J. 2016.Protocol invariance and the timing of decisions in dynamic games. *CEPR working paper* **DP11447**.

[6] Haltiwanger, J. and Harrington, J. 1991. The impact of cyclical demand movements on collusive behaviour. *RAND Journal of Economics* **22**: 89-106.

[7] Harrington Jr. J. E., Zhao W. 2012. Signaling and tacit collusion in an infinitely repeated Prisoners' Dilemma, *Mathematical Social Sciences* **64**: 277-289.

[8] Mailath, G. J. and Samuelson, L. 2006. Repeated Games and Reputations: Long-Run Relationships. Oxford. Oxford University Press.

[9] Olley, G.S. and Pakes,A. 1996. The dynamics of productivity in the telecommunications equipment industry, *Econometrica* 64: 1263-1298.

[10] Pakes, A. and McGuire, P. 2001. Stochastic algorithms, symmetric Markov perfect equilibrium, and the 'curse' of dimensionality, *Econometrica* **69**: 1261-1282.

[11] Puzzello, D. 2008. Tie-breaking rules and divisibility in experimental duopoly markets. *Journal of Economic Behavior and Organization* **67**: 164-179.

[12] Rotemberg, J. and Saloner, G. 1986. A supergame-theoretic model of price wars during booms. *American Economic Review* **76**: 390-407.

[13] Suetens, S. 2008. Does R&D cooperation facilitate price collusion? An experiment. *Journal of Economic Behavior and Organization* **66**: 822–836.

Appendix

Proof of Lemma 1

We begin the proof by calculating the payoff in the strategy profile when players play non-cooperatively:

$$V_n = \begin{pmatrix} V_n^1 \\ V_n^2 \end{pmatrix} = \begin{pmatrix} d_1 + \delta p_1 V_n \\ d_2 + \delta p_2 V_n \end{pmatrix},$$

or in vectorial form:

$$V_n = \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} + \delta \Pi V_n.$$

This equation gives:

$$V_n = (\mathbb{I} - \delta\Pi)^{-1} \begin{pmatrix} d_1 \\ d_2 \end{pmatrix},$$

where $\mathbb{I}$ is an identity matrix of size 2×2. We denote $(\mathbb{I} - \delta\Pi)^{-1}$ by $\widetilde{\Pi}$ and obtain the result.

Second, we calculate the discounted payoff when players play cooperatively, i.e. they play C in any stage of the game:

$$V_c = \begin{pmatrix} V_c^1 \\ V_c^2 \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} + \delta \begin{pmatrix} \pi_1 & 1-\pi_1 \\ \pi_2 & 1-\pi_2 \end{pmatrix} \begin{pmatrix} V_c^1 \\ V_c^2 \end{pmatrix}.$$

Rewriting this equation in vectorial form, we obtain equation (4).

Third, we calculate the discounted payoff of a player when both players play C in state 1 and D in state 2 in any period of the game:

$$V_{sc} = \begin{pmatrix} V_{sc}^1 \\ V_{sc}^2 \end{pmatrix} = \begin{pmatrix} a_1 \\ d_2 \end{pmatrix} + \delta \begin{pmatrix} \pi_1 & 1-\pi_1 \\ \pi_2 & 1-\pi_2 \end{pmatrix} \begin{pmatrix} V_{sc}^1 \\ V_{sc}^2 \end{pmatrix}.$$

Rewriting this equation in a vectorial form, we obtain equation (5).

Finally, we compute the expected payoff of the deviating player j: $V_{d,j}(\delta)$ which is:

$$V_{d,j}(\delta) = \begin{pmatrix} V_{d,j}^1(\delta) \\ V_{d,j}^2(\delta) \end{pmatrix}$$

where $V_{d,j}^s(\delta)$ is the payoff of the player in the subgame starting from state s. If the subgame starts from state 1, player j gets

$$V_{d,j}^1(\delta) = a_1 + \delta(\pi_1 V_{d,j}^1(\delta) + (1-\pi_1)V_{d,j}^2(\delta)),$$

and if the subgame starts from state 2, player j deviates and gets b_2. Then she will be punished by playing (D, D) in any state from the next stage until infinity. Her total payoff will be

$$V_{d,j}^2(\delta) = b_2 + \delta p_2 V_n(\delta).$$

From these two equations we obtain

$$V_{d,j}(\delta) = \begin{pmatrix} \frac{1}{1-\delta\pi_1}[a_1 + \delta(1-\pi_1)(b_2 + \delta p_2 V_n(\delta))] \\ b_2 + \delta p_2 V_n(\delta) \end{pmatrix}.$$

Proof of Proposition 1

Consider first the non-cooperative strategy profile. The fact that the strategy profile σ_n is a SPNE for every δ is true because this profile prescribes players to choose the Nash equilibrium strategies in any state.

Consider next the cooperative strategy profile. The player's payoff is given by Lemma 2. We need to prove that, for

$$\delta \geqslant \frac{b_2 - a_2}{p_2(V_c - V_n)},$$

σ_c is a subgame perfect equilibrium. If the deviation is observed, in the next stage of the game the deviating player is punished by getting the Nash equilibrium payoff in any state (d_1 in state 1 and d_2 in state 2). The strategy profile σ_c is a subgame perfect equilibrium if there is no gain from deviation in any state. The player does not player deviates in state 1 if

$$a_1 + \delta p_1 V_c \geqslant b_1 + \delta p_1 V_n,$$

and in state 2 if

$$a_2 + \delta p_2 V_c \geqslant b_2 + \delta p_2 V_n.$$

Taking into account Assumption 1, two inequalities prove the second part of the proposition.

Finally, consider the semi-cooperative strategy profile. The player's payoff is given by

Lemma 3. We need to prove that, if

$$\delta \geqslant \frac{b_1 - a_1}{p_1(V_{sc} - V_n)},$$

then σ_{sc} is a subgame perfect equilibrium. If the deviation is observed, in the next stage of the game the deviating player is punished by getting the Nash equilibrium payoff in any state (d_1 in state 1 and d_2 in state 2). Clearly, deviation in state 2 is not profitable. Consider state 1. Strategy profile σ_{sc} is a subgame perfect Nash equilibrium if there is no gain from deviation in state 1:

$$a_1 + \delta p_1 V_{sc} \geqslant b_1 + \delta p_1 V_n.$$

This gives the result of Proposition 1.

Proof of Proposition 2

Initial state $s = 1$**. Strategy profile** $(q_1, r_1) = (1, 0)$

In this case, a player of type P colludes in any state, whereas an M player deviates in all states.

The initial state is $s = 1$. Begin from the player type P. If she does not deviate from $(1, 0)$, she gets

$$\alpha_1 \left(a_1 + \delta_P p_1 V_c(\delta_P)\right) + (1 - \alpha_1)\left(c_1 + \delta_P p_1 V_n(\delta_P)\right). \tag{13}$$

If she deviates from profile $(1, 0)$ ($q_1 = 0$), she gets:

$$\alpha_1(b_1 + \delta_P p_1 V_n(\delta_P)) + (1 - \alpha_1)(d_1 + \delta_P p_1 V_n(\delta_P)) \tag{14}$$

The deviation is not profitable if (13) is larger or equal to (14) taking into account $\delta_P \geqslant \delta_2^*$.

Now consider the player of type M. Her payoff in profile $(1, 0)$ is

$$\alpha_1(b_1 + \delta_M p_1 V_n(\delta_M)) + (1 - \alpha_1)(d_1 + \delta_M p_1 V_n(\delta_M)). \tag{15}$$

If she deviates from profile $(1, 0)$ (playing $r_1 = 1$), then she gets:

$$\alpha_1(a_1 + \delta_M p_1 V_d(\delta_M)) + (1 - \alpha_1)(c_1 + \delta_M p_1 V_n(\delta_M)) \tag{16}$$

where $V_d(\delta_M)$ is the payoff of type M player when she cooperates in $s = 1$ and deviates in state $s = 2$ (which is profitable to her according to her discount factor).

The deviation is not profitable if (15) is larger or equal than (16), taking into account inequality $\delta_1^* \leqslant \delta_M \leqslant \delta_2^*$ from Proposition 1.

If the game starts in state $s = 1$, then the strategy profile $(1, 0)$ is a PMPBE when one of the following systems has a solution:

$$\begin{cases} \alpha_1 \geqslant \dfrac{d_1 - c_1}{d_1 - c_1 + a_1 - b_1 + \delta_P p_1(V_c(\delta_P) - V_n(\delta_P))} \\ d_1 - c_1 + a_1 - b_1 + \delta_M p_1(V_d(\delta_M) - V_n(\delta_M)) \leqslant 0 \end{cases}$$

or

$$\begin{cases} \alpha_1 \geqslant \dfrac{d_1 - c_1}{d_1 - c_1 + a_1 - b_1 + \delta_P p_1(V_c(\delta_P) - V_n(\delta_P))} \\ d_1 - c_1 + a_1 - b_1 + \delta_M p_1(V_d(\delta_M) - V_n(\delta_M)) > 0 \\ \alpha_1 \leqslant \dfrac{d_1 - c_1}{d_1 - c_1 + a_1 - b_1 + \delta_M p_1(V_d(\delta_M) - V_n(\delta_M))} \end{cases}$$

Now we need to verify the sign of:

$$d_1 - c_1 + a_1 - b_1 + \delta_M p_1(V_d(\delta_M) - V_n(\delta_M)) > 0. \tag{17}$$

First, consider the difference $V_{sc}(\delta_M) - V_d(\delta_M)$:

$$V_{sc}^1(\delta_M) - V_d^1(\delta_M) = \delta_M \pi_1 (V_{sc}^1(\delta_M) - V_d^1(\delta_M)) + \\ \delta_M (1 - \pi_1)(d_2 - b_2 + \delta_M p_2(V_{sc}(\delta_M) - V_d(\delta_M)),$$

and

$$V_{sc}^1(\delta_M) - V_d^1(\delta_M) = \frac{\delta_M(1 - \pi_1)(d_2 - b_2 + \delta_M p_2(V_{sc}(\delta_M) - V_n(\delta_M))}{1 - \delta_M \pi_1}.$$

Taking into account that

$$\delta_M = \frac{b_2 - a_2}{p_2(V_c(\delta_2^*) - V_n(\delta_2^*))} < \frac{b_2 - a_2}{p_2(V_{sc}(\delta_2^*) - V_n(\delta_2^*))} = \delta_2^*,$$

we prove

$$d_2 - b_2 + \delta_M p_2(V_{sc}(\delta_M) - V_n(\delta_M)) < 0.$$

Therefore, we may state that $V_{sc}^1(\delta_M) - V_d^1(\delta_M) < 0$.

We then calculate the difference:

$$V_{sc}^2(\delta_M) - V_d^2(\delta_M) = d_2 - b_2 + \delta_M p_2 (V_{sc}(\delta_M) - V_n(\delta_M)),$$

which is negative. Taking into account that $V_{sc}(\delta_M) < V_n(\delta_M)$ and the definition of δ_M, the inequality (17) is true.

Simplifying the systems and considering $\delta_P \geqslant \delta_2^*$, we obtain the condition:

$$\alpha_1 \in [A_1, A_2], \tag{18}$$

where

$$A_1 \equiv \frac{d_1 - c_1}{d_1 - c_1 + a_1 - b_1 + \delta_P p_1 (V_c(\delta_P) - V_n(\delta_P))},$$

and

$$A_2 \equiv \frac{d_1 - c_1}{d_1 - c_1 + a_1 - b_1 + \delta_M p_1 (V_d(\delta_M) - V_n(\delta_M))}.$$

Initial state $s = 1$. Strategy profile $(q_1, r_1) = (1, 1)$

Begin from the player type P. If she does not deviate from strategy $(1, 1)$, she gets:

$$(\alpha_1 + \beta_1)(a_1 + \delta_P p_1 V_{sc}(\delta_P)) + (1 - \alpha_1 - \beta_1)(c_1 + \delta_P p_1 V_n(\delta_P)). \tag{19}$$

If she deviates from $(1, 1)$ $(q_1 = 0)$, she gets:

$$(\alpha_1 + \beta_1)(b_1 + \delta_P p_1 V_n(\delta_P)) + (1 - \alpha_1 - \beta_1)(d_1 + \delta_P p_1 V_n(\delta_P)). \tag{20}$$

The deviation is not profitable if (19) is larger or equal to (20), taking into account $\delta_P \geqslant \delta_2^*$ from Proposition 1.

Consider next a player of type M. Her payoff in profile $(1, 1)$ is

$$(\alpha_1 + \beta_1)(a_1 + \delta_M p_1 V_{sc}(\delta_M)) + (1 - \alpha_1 - \beta_1)(c_1 + \delta_M p_1 V_n(\delta_M)). \tag{21}$$

If she deviates from profile $(1, 1)$ $(r_1 = 0)$ she gets:

$$(\alpha_1 + \beta_1)(b_1 + \delta_M p_1 V_n(\delta_M)) + (1 - \alpha_1 - \beta_1)(d_1 + \delta_M p_1 V_n(\delta_M)). \tag{22}$$

The deviation is not profitable if payoff (21) is larger or equal to (22), taking into account

$\delta_1^* \leqslant \delta_M \leqslant \delta_2^*$.

If the game starts in state 1, the strategy profile $(1,1)$ is a PMPBE if the following system has a solution:

$$\begin{cases} (\alpha_1+\beta_1)\left[a_1-b_1-c_1+d_1+\delta_P p_1(V_{sc}(\delta_P)-V_n(\delta_P))\right] \geqslant d_1-c_1 \\ (\alpha_1+\beta_1)\left[a_1-b_1-c_1+d_1+\delta_M p_1(V_{sc}(\delta_M)-V_n(\delta_M))\right] \geqslant d_1-c_1 \end{cases}.$$

Since $\delta_M < \delta_P$, the system is equivalent to the following inequality:

$$\alpha_1+\beta_1 \geqslant \frac{d_1-c_1}{d_1-c_1+a_1-b_1+\delta_M p_1(V_{sc}(\delta_M)-V_n(\delta_M))}. \tag{23}$$

Initial state $s=1$. Strategy profile (q_1, r_1), $q_1 \in (0,1)$, $r_1 \in (0,1)$

With this strategy profile, a player of type P colludes in any state with probability q_1, while a player of type M colludes in state 1 with probability r_1.

Consider first the case when the game starts from state 1. If a player P does not deviate from profile (q_1, r_1), she gets

$$\begin{aligned} &\alpha_1\left\{q_1^2(a_1+\delta_P p_1 V_{sc}(\delta_P))+(1-q_1)q_1(c_1+\delta_P p_1 V_n(\delta_P))\right.\\ &\left.+(1-q_1)q_1(b_1+\delta_P p_1 V_n(\delta_P))+(1-q_1)^2(d_1+\delta_P p_1 V_n(\delta_P))\right\}\\ &+\beta_1\left\{q_1 r_1(a_1+\delta_P p_1 V_{sc}(\delta_P))+(1-q_1)r_1(b_1+\delta_P p_1 V_n(\delta_P))\right.\\ &\left.+(1-r_1)q_1(c_1+\delta_P p_1 V_n(\delta_P))+(1-q_1)(1-r_1)(d_1+\delta_P p_1 V_n(\delta_P))\right\}\\ &+(1-\alpha_1-\beta_1)\left\{q_1(c_1+\delta_P p_1 V_n(\delta_P))+(1-q_1)(d_1+\delta_P p_1 V_n(\delta_P))\right\}. \end{aligned} \tag{24}$$

If she deviates from profile (q_1, r_1) to pure strategy $q_1=0$, she gets:

$$\begin{aligned} &\alpha_1\left\{q_1(b_1+\delta_P p_1 V_n(\delta_P))+(1-q_1)(d_1+\delta_P p_1 V_n(\delta_P))\right\}\\ &+\beta_1\left\{r_1(b_1+\delta_P p_1 V_n(\delta_P))+(1-r_1)(d_1+\delta_P p_1 V_n(\delta_P))\right\}\\ &+(1-\alpha_1-\beta_1)(d_1+\delta_P p_1 V_n(\delta_P)). \end{aligned} \tag{25}$$

If she deviates from profile (q_1, r_1) to pure strategy $q_1 = 1$, then she gets:

$$\begin{aligned}&\alpha_1 \{q_1(a_1 + \delta_P p_1 V_{sc}(\delta_P)) + (1 - q_1)(c_1 + \delta_P p_1 V_n(\delta_P))\} \\ &+\beta_1 \{r_1(a_1 + \delta_P p_1 V_{sc}(\delta_P)) + (1 - r_1)(c_1 + \delta_P p_1 V_n(\delta_P))\} \\ &+(1 - \alpha_1 - \beta_1)(c_1 + \delta_P p_1 V_n(\delta_P)). \qquad (26)\end{aligned}$$

The deviation is not profitable if payoff (24) is larger or equal to (25) and (26) taking into account $\delta_P \geqslant \delta_2^*$. It is true if

$$\alpha_1 q_1 + \beta_1 r_1 = \frac{d_1 - c_1}{d_1 - c_1 + a_1 - b_1 + \delta_P p_1 (V_{sc}(\delta_P) - V_n(\delta_P))}. \qquad (27)$$

If a player M does not deviate from profile (q_1, r_1), she gets

$$\begin{aligned}&\alpha_1 \{q_1 r_1(a_1 + \delta_M p_1 V_{sc}(\delta_M)) + (1 - q_1) r_1 (c_1 + \delta_M p_1 V_n(\delta_M)) \\ &+ (1 - r_1) q_1 (b_1 + \delta_M p_1 V_n(\delta_M)) + (1 - q_1)(1 - r_1)(d_1 + \delta_M p_1 V_n(\delta_M))\} \\ &+\beta_1 \left\{r_1^2(a_1 + \delta_M p_1 V_{sc}(\delta_M)) + (1 - r_1) r_1 (c_1 + \delta_M p_1 V_n(\delta_M)) \right. \\ &\left. + (1 - r_1) r_1 (b_1 + \delta_M p_1 V_n(\delta_M)) + (1 - r_1)^2 (d_1 + \delta_M p_1 V_n(\delta_M))\right\} \\ &+(1 - \alpha_1 - \beta_1) \{r_1(c_1 + \delta_M p_1 V_n(\delta_M)) + (1 - r_1)(d_1 + \delta_M p_1 V_n(\delta_M))\}. \qquad (28)\end{aligned}$$

If she deviates from profile (q_1, r_1) to pure strategy $r_1 = 0$, then she gets:

$$\begin{aligned}&\alpha_1 \{q_1(b_1 + \delta_M p_1 V_n(\delta_M)) + (1 - q_1)(d_1 + \delta_M p_1 V_n(\delta_M))\} \\ &+\beta_1 \{r_1(b_1 + \delta_M p_1 V_n(\delta_M)) + (1 - r_1)(d_1 + \delta_M p_1 V_n(\delta_M))\} \\ &+(1 - \alpha_1 - \beta_1)(d_1 + \delta_M p_1 V_n(\delta_M)). \qquad (29)\end{aligned}$$

If she deviates from profile (q_1, r_1) to pure strategy $r_1 = 1$, then she gets:

$$\begin{aligned}&\alpha_1 \{q_1(a_1 + \delta_M p_1 V_{sc}(\delta_M)) + (1 - q_1)(c_1 + \delta_M p_1 V_n(\delta_M))\} \\ &+\beta_1 \{r_1(a_1 + \delta_M p_1 V_{sc}(\delta_M)) + (1 - r_1)(c_1 + \delta_M p_1 V_n(\delta_M))\} \\ &+(1 - \alpha_1 - \beta_1)(c_1 + \delta_M p_1 V_n(\delta_M)). \qquad (30)\end{aligned}$$

The deviation is not profitable if payoff (28) is larger or equal to (29) and (30) taking into account $\delta_M \geqslant \delta_1^*$. It is true if

$$\alpha_1 q_1 + \beta_1 r_1 = \frac{d_1 - c_1}{d_1 - c_1 + a_1 - b_1 + \delta_M p_1(V_{sc}(\delta_M) - V_n(\delta_M))}. \tag{31}$$

The strategy profile (q_1, r_1), $q_1 \in (0,1)$, $r_1 \in (0,1)$, is a PMPBE if both equations (27) and (31) hold. Taking into account that $\delta_M < \delta_P$ we obtain that both equations hold if $d_1 = c_1$ and $\alpha_1 = \beta_1 = 0$.

Initial state $s = 2$. Strategy profile $(q_1) = (1)$

If the game starts in state 2, the strategy profile (1) is a PMPBE if the following inequality has a solution:

$$\alpha_1 \left[a_2 - b_2 - c_2 + d_2 + \delta_P p_2(V_c(\delta_P) - V_n(\delta_P))\right] \geqslant d_2 - c_2.$$

Since $\delta_P \geqslant \delta_2^*$, then $\delta_P p_2(V_c(\delta_P) - V_n(\delta_P)) \geqslant b_2 - a_2$, so that:

$$\alpha_1 \geqslant \frac{d_2 - c_2}{a_2 - b_2 - c_2 + d_2 + \delta_P p_2(V_c(\delta_P) - V_n(\delta_P))}. \tag{32}$$

Initial state $s = 2$. Strategy profile (q_1), $q_1 \in (0,1)$

If a player P does not deviate from profile (q_1), she gets

$$\begin{aligned} &\alpha_1 \left\{q_1^2(a_2 + \delta_P p_2 V_c(\delta_P)) + (1-q_1)q_1(b_2 + \delta_P p_2 V_n(\delta_P)) \right.\\ &\left. + (1-q_1)q_1(c_2 + \delta_P p_2 V_n(\delta_P)) + (1-q_1)^2(d_2 + \delta_P p_2 V_n(\delta_P))\right\} \\ &+ (1-\alpha_1)\left\{q_1(c_2 + \delta_P p_2 V_n(\delta_P)) + (1-q_1)(d_2 + \delta_P p_2 V_n(\delta_P))\right\}. \end{aligned} \tag{33}$$

If she deviates from profile (q_1, r_1) to pure strategy $q_1 = 0$, she gets:

$$\alpha_1 \left\{q_1(b_2 + \delta_P p_2 V_n(\delta_P)) + (1-q_1)(d_2 + \delta_P p_2 V_n(\delta_P))\right\} + (1-\alpha_1)(d_2 + \delta_P p_2 V_n(\delta_P)). \tag{34}$$

If she deviates from profile (q_1, r_1) to pure strategy $q_1 = 1$, then she gets:

$$\alpha_1 \left\{q_1(a_2 + \delta_P p_2 V_c(\delta_P)) + (1-q_1)(c_2 + \delta_P p_2 V_n(\delta_P))\right\} + (1-\alpha_1)(c_2 + \delta_P p_2 V_n(\delta_P)). \tag{35}$$

The deviation is not profitable if payoff (33) is larger or equal to (34) and (35) taking into account $\delta_P \geqslant \delta_2^*$. It is true if

$$\alpha_1 q_1 = \frac{d_2 - c_2}{d_2 - c_2 + a_2 - b_2 + \delta_P p_2 (V_c(\delta_P) - V_n(\delta_P))}. \tag{36}$$

Therefore, strategy

$$q_1 = q^* \equiv \frac{d_2 - c_2}{\alpha_1 (d_2 - c_2 + a_2 - b_2 + \delta_P p_2 (V_c(\delta_P) - V_n(\delta_P)))}$$

is an optimal strategy if

$$\alpha_1 \geqslant \frac{d_2 - c_2}{d_2 - c_2 + a_2 - b_2 + \delta_P p_2 (V_c(\delta_P) - V_n(\delta_P))}.$$

Proof of Corollary 1

The payoff of a P player in profile $(1, 0)$ is

$$\alpha_1 (a_1 + \delta_P p_1 V_c(\delta_P)) + (1 - \alpha_1)(c_1 + \delta_P p_1 V_n(\delta_P)),$$

and in profile $(1, 1)$ is

$$(\alpha_1 + \beta_1)(a_1 + \delta_P p_1 V_{sc}(\delta_P)) + (1 - \alpha_1 - \beta_1)(c_1 + \delta_P p_1 V_n(\delta_P)).$$

The payoff of a P player in profile $(1, 1)$ is not less than his payoff in profile $(1, 0)$ if

$$\beta_1 (c_1 - a_1 - \delta_P p_1 (V_{sc}(\delta_P) - V_n(\delta_P))) + \alpha_1 \delta_P p_1 (V_c(\delta_P) - V_{sc}(\delta_P)) \leqslant 0,$$

or

$$\frac{\beta_1}{\alpha_1} \geqslant \frac{\delta_P p_1 (V_c(\delta_P) - V_{sc}(\delta_P))}{a_1 - c_1 + \delta_P p_1 (V_{sc}(\delta_P) - V_n(\delta_P))}.$$

The payoff of an M player in profile $(1, 0)$ is

$$\alpha_1 (b_1 + \delta_M p_1 V_n(\delta_M)) + (1 - \alpha_1)(d_1 + \delta_M p_1 V_n(\delta_M))$$

and in profile $(1,1)$ is

$$(\alpha_1+\beta_1)(a_1+\delta_M p_1 V_{sc}(\delta_M))+(1-\alpha_1-\beta_1)(c_1+\delta_M p_1 V_n(\delta_M)).$$

The payoff of an M player in profile $(1,1)$ is not less than his payoff in profile $(1,0)$ if

$$\alpha_1(b_1-a_1+c_1-d_1+\delta_M p_1(V_n(\delta_M)-V_{sc}(\delta_M)))+\beta_1(c_1-a_1+\delta_M p_1(V_n(\delta_M)-V_{sc}(\delta_M))) + d_1-c_1 \leqslant 0.$$

or

$$(\alpha_1+\beta_1)(d_1-b_1+a_1-c_1+\delta_M p_1(V_{sc}(\delta_M)-V_n(\delta_M))\geqslant \beta_1(d_1-b_1)+(d_1-c_1). \quad (37)$$

Taking into account that $\alpha_1+\beta_1 \geqslant A_3$, we may state that

$$(\alpha_1+\beta_1)(d_1-b_1+a_1-c_1+\delta_M p_1(V_{sc}(\delta_M)-V_n(\delta_M))\geqslant d_1-c_1.$$

The latter inequality guarantees that (37) is satisfied because $d_1-b_1<0$.

Now consider the initial state $s=2$. The payoff of type P player in equilibrium $(q_1^2)=1$ is

$$\alpha_1(a_2+\delta_P p_2 V_c(\delta_P))+(1-\alpha_1)(c_2+\delta_P p_2 V_n(\delta_P)).$$

The payoff of type P player in equilibrium $(q_1^2)=q$ given by $q^*=A_4/\alpha_1$ (see Proposition 2) is

$$\alpha_1 qq(a_2-b_2+d_2-c_2+\delta_P p_2(V_c(\delta_P)-V_n(\delta_P))+\alpha_1 q(b_2-d_2)+q(c_2-d_2)+d_2+\delta_P p_2 V_n(\delta_P) = d_2+\alpha_1 q(b_2-d_2)+\delta_P p_2 V_n(\delta_P).$$

The payoff of type P player in profile $(q_1^2)=1$ is not less than his payoff in profile $(q_1^2)=q$ if

$$\alpha_1(a_2-c_2+\delta_P p_2(V_c(\delta_P)-V_n(\delta_P))) \geqslant (d_2-c_2)\left[1+\frac{b_2-d_2}{d_2-c_2+a_2-b_2+\delta_P p_2(V_c(\delta_P)-V_n(\delta_P))}\right],$$

which is always true for any $\alpha_1 \geqslant A_4$.

The payoff of player M in profile $(q_1^2) = 1$ is

$$\alpha_1(b_2 + \delta_M p_2 V_n(\delta_M)) + (1 - \alpha_1)(d_2 + \delta_M p_2 V_n(\delta_M)),$$

and in profile $(q_1^2) = q$ is

$$\alpha_1(q(b_2 + \delta_M p_2 V_n(\delta_M)) + (1 - q)(d_2 + \delta_M p_2 V_n(\delta_M))) + (1 - \alpha_1)(d_2 + \delta_M p_2 V_n(\delta_M)).$$

The payoff of type M player in profile $(q_1^2) = 1$ is not less than his payoff in profile $(q_1^2) = q$ if

$$\alpha_1(b_2 - qb_2 - (1 - q)d_2) \geqslant 0,$$

which is always true because $b_2 > d_2$.

Proof of Proposition 3

Consider the payoffs of the players of types P and M as functions of parameter α_1. By Proposition 2, there are three equilibria:

1. Equilibrium $(q_1^1, r_1^1) = (1, 0)$: the payoff of the player of type P is

   $$\alpha_1(a_1 + \delta_P p_1 V_c(\delta_P)) + (1 - \alpha_1)(c_1 + \delta_P p_1 V_n(\delta_P)).$$

   It is a linear function of α_1 with coefficient $a_1 - c_1 + \delta_P p_1 (V_c(\delta_P) - V_n(\delta_P))$ which is positive because $a_1 > c_1$ and $V_c(\delta) > V_n(\delta)$ for any $\delta \in (0, 1)$.

   The payoff of the player of type M is

   $$\alpha_1(b_1 + \delta_M p_1 V_n(\delta_M)) + (1 - \alpha_1)(d_1 + \delta_M p_1 V_n(\delta_M)).$$

   It is also a linear function of α_1 with coefficient $b_1 - d_1$ which is positive for any $\delta \in (0, 1)$.

2. Equilibrium $(q_1^1, r_1^1) = (1, 1)$: we begin with the player of type P. Her payoff is

   $$(\alpha_1 + \beta_1)(a_1 + \delta_P p_1 V_{sc}(\delta_P)) + (1 - \alpha_1 - \beta_1)(c_1 + \delta_P p_1 V_n(\delta_P)).$$

   It is a a linear function of α_1 with coefficient $a_1 - c_1 + \delta_P p_1 (V_{sc}(\delta_P) - V_n(\delta_P))$ which is positive because $a_1 > c_1$ and $V_{sc}(\delta) > V_n(\delta)$ for any $\delta \in (0, 1)$.

Then, the payoff of the player of type M is

$$(\alpha_1 + \beta_1)(a_1 + \delta_M p_1 V_{sc}(\delta_M)) + (1 - \alpha_1 - \beta_1)(c_1 + \delta_M p_1 V_n(\delta_M)).$$

This is a linear function of α_1 with coefficient $a_1 - c_1 + \delta_M p_1 (V_{sc}(\delta_M) - V_n(\delta_M))$ which is positive because $a_1 > c_1$ and $V_{sc}(\delta) > V_n(\delta)$ for any $\delta \in (0,1)$.

The derivatives of the payoffs of the P and M players with respect to β_1 equal the corresponding derivatives subject to α_1. Therefore, the payoffs are also increasing functions of β_1.

3. Equilibrium $(q_1) = (1)$ in initial state $s = 2$: the payoff of the player of type P is

$$\alpha_1(a_2 + \delta_P p_2 V_c(\delta_P)) + (1 - \alpha_1)(c_2 + \delta_P p_2 V_n(\delta_P))$$

It is a linear function of α_1 with coefficient $a_2 - c_2 + \delta_P p_2 (V_c(\delta_P) - V_n(\delta_P))$ which is positive because $a_2 > c_2$ and $V_c(\delta) > V_n(\delta)$ for any $\delta \in (0,1)$.

4. Equilibrium $(q_1) = q^*$ in initial state $s = 2$. , where

$$q^* = \frac{d_2 - c_2}{\alpha_1(d_2 - c_2 + a_2 - b_2 + \delta_P p_2 (V_c(\delta_P) - V_n(\delta_P)))}.$$

The payoff of the player of type P is

$$\begin{aligned} &\alpha_1 \{q_1(a_2 + \delta_P p_2 V_c(\delta_P)) + (1 - q_1)q_1(b_2 + \delta_P p_2 V_n(\delta_P)) \\ &+ (1 - q_1)q_1(c_2 + \delta_P p_2 V_n(\delta_P)) + (1 - q_1)(d_2 + \delta_P p_2 V_n(\delta_P))\} \\ &+(1 - \alpha_1)\{q_1(c_2 + \delta_P p_2 V_n(\delta_P)) + (1 - q_1)(d_2 + \delta_P p_2 V_n(\delta_P))\}. \end{aligned}$$

Substituting q^* into the payoff and finding the derivative s.t. α_1, we obtain that the derivative equals

$$\frac{A_4^2}{2\alpha_1^2}(b_2 + c_2 + 2\delta_P p_2 V_n(\delta_P)) + \frac{A_4}{2\alpha_1^2}(d_2 - c_2),$$

which is positive because $d_2 > c_2$.

Proof of Proposition 4

Consider the strategy profile $(q_1; r_1) = (0, 0)$. A player P obtains the following payoff if she does not deviate from $(0, 0)$:

$$\begin{pmatrix} d_1 + \delta_P p_1 V_n(\delta_P) \\ d_2 + \delta_P p_2 V_n(\delta_P) \end{pmatrix}. \tag{38}$$

If she deviates from profile $(0, 0)$ $(q_1 = 1)$, she gets:

$$\begin{pmatrix} c_1 + \delta_P p_1 V_n(\delta_P) \\ c_2 + \delta_P p_2 V_n(\delta_P) \end{pmatrix}. \tag{39}$$

Note that (38) is always greater or equal to (39), since $d_i \geqslant c_i$ for any $i \in \{1, 2\}$. A deviation of a type M cannot be profitable either. Therefore, the strategy profile $(q_1, r_1) = (0, 0)$ is a PMPBE, but in this case neither cooperative nor semi-cooperative strategy profile is realised in the collusion phase of the game.